PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 9/02, 15/00, C07H 21/02, 21/04

A1

(11) International Publication Number: WO 00/05350

(43) International Publication Date: 3 February 2000 (03.02.00)

(21) International Application Number: PCT/US99/16746

(22) International Filing Date: 23 July 1999 (23.07.99)

(30) Priority Data:
60/094,012 24 July 1998 (24.07.98) US

(71) Applicant *(for all designated States except US)*: WASHINGTON STATE UNIVERSITY RESEARCH FOUNDATION [US/US]; N.E. 1615 Eastgate Boulevard, Pullman, WA 99164–1802 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LEWIS, Norman, G. [CA/US]; 1710 Upper Drive, Pullman, WA 99163 (US). KASAHARA, Hiroyuki [JP/US]; Apartment 8, 215 N.W. Anthony Street, Pullman, WA 99163 (US). GANG, David, R. [US/US]; 2336 Stone Road, Ann Arbor, MI 48108 (US). DAVIN, Laurence, B. [FR/US]; 1710 Upper Drive, Pullman, WA 99163 (US).

(74) Agent: McGURL, Barry, F.; Christensen O'Connor Johnson & Kindness PLLC Suite 2800, 1420 Fifth Avenue, Seattle, WA 98101 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: RECOMBINANT DEHYDRODICONIFERYL ALCOHOL BENZYLIC ETHER REDUCTASE, AND METHODS OF USE

(57) Abstract

Isolated DNA sequences are provided which code for the expression of dehydrodiconiferyl alcohol benzylic ether reductase. In other aspects, replicable recombinant cloning vehicles are provided which code for dehydrodiconiferyl alcohol benzylic ether reductase or for a base sequence sufficiently complementary to at least a portion of dehydrodiconiferyl alcohol benzylic ether reductase DNA or RNA to enable hybridization therewith. In yet other aspects, modified host cells are provided that have been transformed, transfected, infected and/or injected with a recombinant cloning vehicle and/or DNA sequence encoding dehydrodiconiferyl alcohol benzylic ether reductase. Thus, methods are provided for the recombinant expression of dehydrodiconiferyl alcohol benzylic ether reductase in host cells, such as plant cells.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

RECOMBINANT DEHYDRODICONIFERYL ALCOHOL BENZYLIC ETHER REDUCTASE, AND METHODS OF USE

Related Applications

The present application claims benefit of priority from United States provisional patent application serial number 60/094,012, filed July 24, 1998.

Field of the Invention

The present invention relates to isolated dehydrodiconiferyl alcohol benzylic ether reductase proteins, to nucleic acid sequences which code for dehydrodiconiferyl alcohol benzylic ether reductase proteins, and to vectors containing the sequences, host cells containing the sequences and methods of producing recombinant dehydrodiconiferyl alcohol benzylic ether reductase proteins and their mutants.

Background of the Invention

Dehydrodiconiferyl alcohol (DDC) is an 8,5′-linked lignan and one of three major products of the *in vitro* radical coupling of *E*-coniferyl alcohol that is facilitated by peroxidases and laccases, such as those proposed in lignification (Freudenberg, K. *Bull. Soc. Chim. France* , 1748-1753 (1959); Freudenberg, K. & Neish, A.C. *Constitution and Biosynthesis of Lignin* 1-123 (Springer-Verlag, New York, NY, 1968)). DDC is ubiquitous in the plant kingdom, being found in plants as diverse as loblolly pine (*Pinus taeda*) (Nose, M., *et al. Phytochemistry* **39**:71-79 (1995)) and tobacco (*Nicotiana tabacum*) (Binns, A., N., Chen, R., H., Wood, H., N. & Lynn, D., G. *Proc. Natl. Acad. Sci. USA* **84**:980-984 (1987)). Suspension culture cells of loblolly pine have also been shown to contain DDC and its 7′,8′-(allylic bond) reduced derivative, dihydrodehydrodiconiferyl alcohol (DDDC) (Nose, M., *et al.*

Phytochemistry **39**:71-79 (1995)). However, the enzymes involved in their formation and subsequent metabolism have not been previously described.

Summary of the Invention

In accordance with the foregoing, cDNAs encoding dehydrodiconiferyl alcohol benzylic ether reductase from *Pinus taeda* and *Cryptomeria japonica* have been isolated and sequenced, and the corresponding amino acid sequences have been deduced. Accordingly, the present invention relates to isolated DNA sequences which code for the expression of dehydrodiconiferyl alcohol benzylic ether reductase, such as the sequence designated SEQ ID NO:1 which encodes a dehydrodiconiferyl alcohol benzylic ether reductase (SEQ ID NO:2) from *Pinus taeda*, and the sequences designated SEQ ID NO:3 and SEQ ID NO:5 which encode dehydrodiconiferyl alcohol benzylic ether reductases (SEQ ID NO:4 and SEQ ID NO:6, respectively) from *Cryptomeria japonica*. In other aspects, the present invention is directed to replicable recombinant cloning vehicles comprising a nucleic acid sequence, e.g., a DNA sequence which codes for a dehydrodiconiferyl alcohol benzylic ether reductase, or for a base sequence sufficiently complementary to at least a portion of DNA or RNA encoding dehydrodiconiferyl alcohol benzylic ether reductase to enable hybridization therewith (e.g., antisense RNA or fragments of DNA complementary to a portion of DNA or RNA molecules encoding dehydrodiconiferyl alcohol benzylic ether reductase which are useful as polymerase chain reaction primers or as probes for any of the foregoing reductases or related genes). In yet other aspects of the invention, modified host cells are provided that have been transformed, transfected, infected and/or injected with a recombinant cloning vehicle and/or DNA sequence of the invention. Thus, the present invention provides for the recombinant expression of dehydrodiconiferyl alcohol benzylic ether reductase, and the inventive concepts may be used to facilitate the production, isolation and purification of significant quantities of recombinant dehydrodiconiferyl alcohol benzylic ether reductase (or of its primary enzyme product) for subsequent use, to obtain expression or enhanced expression of dehydrodiconiferyl alcohol benzylic ether reductase in plants, microorganisms or animals, or may be otherwise employed in an environment where the regulation or expression of dehydrodiconiferyl alcohol benzylic ether reductase is desired for the production of this reductase, or its enzyme products, or derivatives thereof. In another aspect, the present invention relates to methods of enhancing or otherwise modifying the expression of dehydrodiconiferyl alcohol benzylic ether reductase protein in a suitable host cell, such as a plant cell.

Brief Description of the Drawings

The foregoing aspects and many of the attendant advantages of this invention will become more readily appreciated as the same becomes better understood by reference to the following detailed description, when taken in conjunction with the accompanying drawings, wherein:

FIGURE 1A shows HPLC separations of the following standards: Peak 1: Tetrahydrodehydrodiconiferyl alcohol (TDDC), Peak 2: Isodihydrodehydrodiconiferyl alcohol (IDDDC), Peak 3: Dihydrodehydrodiconiferyl alcohol (DDDC), Peak 4: Dehydrodiconiferyl alcohol (DDC).

FIGURE 1B shows an HPLC chromatogram of the reduction of DDC to IDDDC.

FIGURE 1C shows an HPLC chromatogram of the reduction of DDDC to TDDC by the benzylic ether reductase (SEQ ID NO:2).

FIGURE 1D shows the mass spectrum of DDC (substrate).

FIGURE 1E shows the mass spectrum of IDDDC (product of DDC reduction by the benzylic ether reductase (SEQ ID NO:2)).

FIGURE 1F shows the mass spectrum of TDDC (product of DDDC reduction by the benzylic ether reductase (SEQ ID NO:2)).

Detailed Description of the Preferred Embodiment

As used herein, the terms "amino acid" and "amino acids" refer to all naturally occurring L-α-amino acids or their residues. The amino acids are identified by either the single-letter or three-letter designations:

Asp	D	aspartic acid	Ile	I	isoleucine
Thr	T	threonine	Leu	L	leucine
Ser	S	serine	Tyr	Y	tyrosine
Glu	E	glutamic acid	Phe	F	phenylalanine
Pro	P	proline	His	H	histidine
Gly	G	glycine	Lys	K	lysine
Ala	A	alanine	Arg	R	arginine
Cys	C	cysteine	Trp	W	tryptophan
Val	V	valine	Gln	Q	glutamine
Met	M	methionine	Asn	N	asparagine

As used herein, the term "nucleotide" means a monomeric unit of DNA or RNA containing a sugar moiety (pentose), a phosphate and a nitrogenous heterocyclic

base. The base is linked to the sugar moiety via the glycosidic carbon (1' carbon of pentose) and that combination of base and sugar is called a nucleoside. The base characterizes the nucleotide with the four bases of DNA being adenine ("A"), guanine ("G"), cytosine ("C") and thymine ("T"). Inosine ("I") is a synthetic base that can be used to substitute for any of the four, naturally-occurring bases (A, C, G or T). The four RNA bases are A,G,C and uracil ("U"). The nucleotide sequences described herein comprise a linear array of nucleotides connected by phosphodiester bonds between the 3' and 5' carbons of adjacent pentoses.

"Oligonucleotide" refers to short length single or double stranded sequences of deoxyribonucleotides linked via phosphodiester bonds. The oligonucleotides are chemically synthesized by known methods and purified, for example, on polyacrylamide gels.

The term "dehydrodiconiferyl alcohol benzylic ether reductase" as used herein means an enzyme that is capable of converting dehydrodiconiferyl alcohol to 7-O-4'-isodihydrodehydrodiconiferyl alcohol, as determined, for example, in the assay described in Example 3 herein.

The terms "alteration", "amino acid sequence alteration", "variant" and "amino acid sequence variant" refer to dehydrodiconiferyl alcohol benzylic ether reductase molecules with some differences in their amino acid sequences as compared to the corresponding native dehydrodiconiferyl alcohol benzylic ether reductase. Ordinarily, the variants will possess at least about 70% homology with the corresponding, native dehydrodiconiferyl alcohol benzylic ether reductase, and preferably they will be at least about 80% homologous with the corresponding, native dehydrodiconiferyl alcohol benzylic ether reductase. The amino acid sequence variants of dehydrodiconiferyl alcohol benzylic ether reductase falling within this invention possess substitutions, deletions, and/or insertions at certain positions. Sequence variants of dehydrodiconiferyl alcohol benzylic ether reductase may be used to attain desired enhanced or reduced enzymatic activity, modified regiochemistry or stereochemistry, or altered substrate utilization or product distribution.

Substitutional dehydrodiconiferyl alcohol benzylic ether reductase variants are those that have at least one amino acid residue in the corresponding native dehydrodiconiferyl alcohol benzylic ether reductase sequence removed and a different amino acid inserted in its place at the same position. The substitutions may be single, where only one amino acid in the molecule has been substituted, or they may be multiple, where two or more amino acids have been substituted in the same molecule.

Substantial changes in the activity of the dehydrodiconiferyl alcohol benzylic ether reductase molecule may be obtained by substituting an amino acid with a side chain that is significantly different in charge and/or structure from that of the native amino acid. This type of substitution would be expected to affect the structure of the polypeptide backbone and/or the charge or hydrophobicity of the molecule in the area of the substitution.

Moderate changes in the activity of the dehydrodiconiferyl alcohol benzylic ether reductase molecule would be expected by substituting an amino acid with a side chain that is similar in charge and/or structure to that of the native molecule. This type of substitution, referred to as a conservative substitution, would not be expected to substantially alter either the structure of the polypeptide backbone or the charge or hydrophobicity of the molecule in the area of the substitution.

Insertional dehydrodiconiferyl alcohol benzylic ether reductase variants are those with one or more amino acids inserted immediately adjacent to an amino acid at a particular position in the native dehydrodiconiferyl alcohol benzylic ether reductase molecule. Immediately adjacent to an amino acid means connected to either the α-carboxy or α-amino functional group of the amino acid. The insertion may be one or more amino acids. Ordinarily, the insertion will consist of one or two conservative amino acids. Amino acids similar in charge and/or structure to the amino acids adjacent to the site of insertion are defined as conservative. Alternatively, this invention includes insertion of an amino acid with a charge and/or structure that is substantially different from the amino acids adjacent to the site of insertion.

Deletional variants are those where one or more amino acids in the native dehydrodiconiferyl alcohol benzylic ether reductase molecule have been removed. Ordinarily, deletional variants will have one or two amino acids deleted in a particular region of the dehydrodiconiferyl alcohol benzylic ether reductase molecule.

The term "antisense" or "antisense RNA" or "antisense nucleic acid" is used herein to mean a nucleic acid molecule that is complementary to all or part of a messenger RNA molecule. Antisense nucleic acid molecules are typically used to inhibit the expression, *in vivo*, of complementary, expressed messenger RNA molecules.

Amino acid sequence variants of dehydrodiconiferyl alcohol benzylic ether reductase may have desirable altered biological activity including, for example, altered reaction kinetics, substrate utilization, product distribution or other characteristics such as regiochemistry and stereochemistry.

The terms "DNA sequence encoding", "DNA encoding" and "nucleic acid encoding" refer to the order or sequence of deoxyribonucleotides along a strand of deoxyribonucleic acid. The order of these deoxyribonucleotides determines the order of amino acids along the translated polypeptide chain. The DNA sequence thus codes for the amino acid sequence.

The terms "vector", "replicable expression vector" and "expression vector" refer to a piece of DNA, usually double-stranded, which may have inserted into it a piece of DNA (insert DNA). The vector is used to transport the insert DNA into a suitable host cell. Once in the host cell, the vector can replicate independently of or coincidentally with the host chromosomal DNA, and several copies of the vector and its inserted DNA may be generated. In addition, expression vectors (including replicable expression vectors) contain the necessary elements that permit translating the insert DNA into a polypeptide. Many molecules of the polypeptide encoded by the insert DNA can thus be rapidly synthesized.

The terms "transformed host cell," "transformed" and "transformation" refer to the introduction of DNA into a cell. The cell is termed a "host cell", and it may be a prokaryotic or a eukaryotic cell. Typical prokaryotic host cells include various strains of *E. coli*. Typical eukaryotic host cells are plant cells, such as maize cells, yeast cells, insect cells or animal cells. The introduced DNA is usually in the form of a vector containing an inserted piece of DNA. The introduced DNA sequence may be from the same species as the host cell or from a different species from the host cell, or it may be a hybrid DNA sequence, containing some foreign DNA and some DNA derived from the host species.

In accordance with the present invention, a cDNA molecule encoding a dehydrodiconiferyl alcohol benzylic ether reductase (SEQ ID NO:1) was isolated from a λ phage-based cDNA library made from RNA extracted from *Pinus taeda* cell suspension culture cells that had been grown in 2,4-D medium. This cDNA library was screened using a 5'-end fragment (SEQ ID NO:7) from a cDNA (PLR-Fi1) encoding pinoresinol-lariciresinol reductase from *Forsythia intermedia*. Twenty positive phage plaques were purified and the cDNA inserts sequenced revealing a putative reductase clone. The β-galactosidase fusion protein encoded by this clone lacked pinoresinol-lariciresinol activity. The cDNA (SEQ ID NO:1) encoding the putative reductase was cloned into the expression plasmid pSBETa, expressed in *E. coli* and the resulting, native protein (SEQ ID NO:2), *i.e.*, lacking a β-galactosidase domain, was purified from a crude *E. coli* extract and assayed for both pinoresinol-

lariciresinol activity and its ability to reduce dehydrodiconiferyl alcohol. The purified *P. taeda* reductase (SEQ ID NO:2) effected the reduction of the benzylic ether bond of dehydrodiconiferyl alcohol to convert it to 7-O-4'-(iso)dihydrodehydrodiconiferylalcohol.

Additionally, two cDNA molecules encoding dehydrodiconiferyl alcohol benzylic ether reductase were isolated from a *Cryptomeria japonica* cDNA library in the following manner. A *C. japonica* cDNA library was screened using 5 ng PCR-amplified *Pinus taeda* dehydrodiconiferyl alcohol benzylic ether reductase cDNA (SEQ ID NO:1). Approximately 300,000 pfu of *C. japonica* amplified cDNA library were screened and yielded twenty positive plaques of which two, pCj-PCBER1 (SEQ ID NO:3) and pCj-PCBER2 (SEQ ID NO:5), were each found to encode a dehydrodiconiferyl alcohol benzylic ether reductase (SEQ ID NO:4 and SEQ ID NO:6, respectively).

The isolation of cDNAs encoding dehydrodiconiferyl alcohol benzylic ether reductase permits the development of an efficient expression system for this functional enzyme, provides useful tools for examining the developmental regulation of lignan biosynthesis and permits the isolation of other dehydrodiconiferyl alcohol benzylic ether reductases. The isolation of dehydrodiconiferyl alcohol benzylic ether reductase cDNAs also permits the transformation of a wide range of organisms in order to enhance or modify lignan biosynthesis.

By way of non-limiting example, the proteins and nucleic acids of the present invention can be utilized to elevate or otherwise alter the levels of lignans in plant species, and in food items incorporating material derived from such genetically altered plants. A nucleic acid sequence encoding dehydrodiconiferyl alcohol benzylic ether reductase, or an antisense nucleic acid fragment complementary to all or part of a nucleic acid sequence encoding dehydrodiconiferyl alcohol benzylic ether reductase, may be introduced, as appropriate, into any plant species for a variety of purposes including, but not limited to: altering or improving the color, texture, durability and pest-resistance of wood tissue, especially heartwood tissue; reducing or altering the formation of lignans and/or lignins in plant species; reducing or altering the lignan/lignin content of plant species utilized in pulp and paper production, thereby making pulp and paper production easier and cheaper; improving the defensive capability of a plant against predators and pathogens by enhancing the production of defensive lignans or lignins; the alteration of other ecological interactions mediated by lignans or lignins; introducing, enhancing or inhibiting the production of

dehydrodiconiferyl alcohol benzylic ether reductases, or their products. A nucleic acid sequence encoding dehydrodiconiferyl alcohol benzylic ether reductase may be introduced into any organism for a variety of purposes including, but not limited to: introducing, enhancing or inhibiting the production of dehydrodiconiferyl alcohol benzylic ether reductase, or the production of its enzymatic product or derivatives thereof.

N-terminal transport sequences well known in the art (see, e.g., von Heijne, G. et al., *Eur. J. Biochem* **180**:535-545 (1989); Stryer, *Biochemistry* W.H. Freeman and Company, New York, NY, p. 769 (1988)) may be employed to direct dehydrodiconiferyl alcohol benzylic ether reductase protein to a variety of cellular or extracellular locations.

Sequence variants of wild-type dehydrodiconiferyl alcohol benzylic ether reductase clones that can be produced by deletions, substitutions, mutations and/or insertions are intended to be within the scope of the invention except insofar as limited by the prior art. Dehydrodiconiferyl alcohol benzylic ether reductase amino acid sequence variants may be constructed by mutating a DNA sequence that encodes wild-type dehydrodiconiferyl alcohol benzylic ether reductase, such as by using techniques commonly referred to as site-directed mutagenesis. Various polymerase chain reaction (PCR) methods now well known in the field, such as a two primer system like the Transformer Site-Directed Mutagenesis kit from Clontech, may be employed for this purpose.

Following denaturation of the target plasmid in this system, two primers are simultaneously annealed to the plasmid; one of these primers contains the desired site-directed mutation, the other contains a mutation at another point in the plasmid resulting in elimination of a restriction site. Second strand synthesis is then carried out, tightly linking these two mutations, and the resulting plasmids are transformed into a *mutS* strain of *E. coli*. Plasmid DNA is isolated from the transformed bacteria, restricted with the relevant restriction enzyme (thereby linearizing the unmutated plasmids), and then retransformed into *E. coli*. This system allows for generation of mutations directly in an expression plasmid, without the necessity of subcloning or generation of single-stranded phagemids. The tight linkage of the two mutations and the subsequent linearization of unmutated plasmids results in high mutation efficiency and allows minimal screening. Following synthesis of the initial restriction site primer, this method requires the use of only one new primer type per mutation site. Rather than prepare each positional mutant separately, a set of "designed degenerate"

oligonucleotide primers can be synthesized in order to introduce all of the desired mutations at a given site simultaneously. Transformants can be screened by sequencing the plasmid DNA through the mutagenized region to identify and sort mutant clones. Each mutant DNA can then be restricted and analyzed by electrophoresis on Mutation Detection Enhancement gel (J.T. Baker) to confirm that no other alterations in the sequence have occurred (by band shift comparison to the unmutagenized control).

The verified mutant duplexes can be cloned into a replicable expression vector, if not already cloned into a vector of this type, and the resulting expression construct used to transform *E. coli*, such as strain *E. coli* BL21(DE3)pLysS, for high level production of the mutant protein, and subsequent purification thereof. The method of FAB-MS mapping can be employed to rapidly check the fidelity of mutant expression. This technique provides for sequencing segments throughout the whole protein and provides the necessary confidence in the sequence assignment. In a mapping experiment of this type, protein is digested with a protease (the choice will depend on the specific region to be modified since this segment is of prime interest and the remaining map should be identical to the map of unmutagenized protein). The set of cleavage fragments is fractionated by microbore HPLC (reversed phase or ion exchange, again depending on the specific region to be modified) to provide several peptides in each fraction, and the molecular weights of the peptides are determined by FAB-MS. The masses are then compared to the molecular weights of peptides expected from the digestion of the predicted sequence, and the correctness of the sequence quickly ascertained. Since this mutagenesis approach to protein modification is directed, sequencing of the altered peptide should not be necessary if the MS agrees with prediction. If necessary to verify a changed residue, CAD-tandem MS/MS can be employed to sequence the peptides of the mixture in question, or the target peptide purified for subtractive Edman degradation or carboxypeptidase Y digestion depending on the location of the modification.

In the design of a particular site directed mutant, it is generally desirable to first make a non-conservative substitution (e.g., Ala for Cys, His or Glu) and determine if activity is greatly impaired as a consequence. The properties of the mutagenized protein are then examined with particular attention to the kinetic parameters of K_m and k_{cat} as sensitive indicators of altered function, from which changes in binding and/or catalysis *per se* may be deduced by comparison to the native enzyme. If the residue is by this means demonstrated to be important by

activity impairment, or knockout, then conservative substitutions can be made, such as Asp for Glu to alter side chain length, Ser for Cys, or Arg for His. For hydrophobic segments, it is largely size that will be altered, although aromatics can also be substituted for alkyl side chains. Changes in the normal product distribution can indicate which step(s) of the reaction sequence have been altered by the mutation.

Other site directed mutagenesis techniques may also be employed with the nucleotide sequences of the invention. For example, restriction endonuclease digestion of DNA followed by ligation may be used to generate dehydrodiconiferyl alcohol benzylic ether reductase deletion variants, as described in Section 15.3 of Sambrook et al. (*Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Press, New York, NY (1989)). A similar strategy may be used to construct insertion variants, as described in Section 15.3 of Sambrook et al., *supra*.

Oligonucleotide-directed mutagenesis may also be employed for preparing substitution variants of this invention. It may also be used to conveniently prepare the deletion and insertion variants of this invention. This technique is well known in the art as described by Adelman et al. (*DNA* **2**:183 (1983)). Generally, oligonucleotides of at least 25 nucleotides in length are used to insert, delete or substitute two or more nucleotides in a dehydrodiconiferyl alcohol benzylic ether reductase gene. An optimal oligonucleotide will have 12 to 15 perfectly matched nucleotides on either side of the nucleotides coding for the mutation. To mutagenize a wild-type dehydrodiconiferyl alcohol benzylic ether reductase, the oligonucleotide is annealed to the single-stranded DNA template molecule under suitable hybridization conditions. A DNA polymerizing enzyme, usually the Klenow fragment of *E. coli* DNA polymerase I, is then added. This enzyme uses the oligonucleotide as a primer to complete the synthesis of the mutation-bearing strand of DNA. Thus, a heteroduplex molecule is formed such that one strand of DNA encodes the wild-type dehydrodiconiferyl alcohol benzylic ether reductase inserted in the vector, and the second strand of DNA encodes the mutated form of dehydrodiconiferyl alcohol benzylic ether reductase inserted into the same vector. This heteroduplex molecule is then transformed into a suitable host cell.

Mutants with more than one amino acid substituted may be generated in one of several ways. If the amino acids are located close together in the polypeptide chain, they may be mutated simultaneously using one oligonucleotide that codes for all of the desired amino acid substitutions. If however, the amino acids are located some distance from each other (separated by more than ten amino acids, for example)

it is more difficult to generate a single oligonucleotide that encodes all of the desired changes. Instead, one of two alternative methods may be employed. In the first method, a separate oligonucleotide is generated for each amino acid to be substituted. The oligonucleotides are then annealed to the single-stranded template DNA simultaneously, and the second strand of DNA that is synthesized from the template will encode all of the desired amino acid substitutions.

An alternative method involves two or more rounds of mutagenesis to produce the desired mutant. The first round is as described for the single mutants: wild-type dehydrodiconiferyl alcohol benzylic ether reductase DNA is used for the template, an oligonucleotide encoding the first desired amino acid substitution(s) is annealed to this template, and the heteroduplex DNA molecule is then generated. The second round of mutagenesis utilizes the mutated DNA produced in the first round of mutagenesis as the template. Thus, this template already contains one or more mutations. The oligonucleotide encoding the additional desired amino acid substitution(s) is then annealed to this template, and the resulting strand of DNA now encodes mutations from both the first and second rounds of mutagenesis. This resultant DNA can be used as a template in a third round of mutagenesis, and so on.

Eukaryotic expression systems may be utilized for dehydrodiconiferyl alcohol benzylic ether reductase production since they are capable of carrying out any required posttranslational modifications and of directing the enzyme to the proper membrane location. A representative eukaryotic expression system for this purpose uses the recombinant baculovirus, *Autographa californica* nuclear polyhedrosis virus (AcNPV; M.D. Summers and G.E. Smith, *A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures* (1986); Luckow et al., *Bio-technology* **6**:47-55 (1987)) for expression of the dehydrodiconiferyl alcohol benzylic ether reductases of the invention. Infection of insect cells (such as cells of the species *Spodoptera frugiperda*) with the recombinant baculoviruses allows for the production of large amounts of the dehydrodiconiferyl alcohol benzylic ether reductase protein. In addition, the baculovirus system has other important advantages for the production of recombinant dehydrodiconiferyl alcohol benzylic ether reductase. For example, baculoviruses do not infect humans and can therefore be safely handled in large quantities. In the baculovirus system, a DNA construct is prepared including a DNA segment encoding dehydrodiconiferyl alcohol benzylic ether reductase and a vector. The vector may comprise the polyhedron gene promoter region of a baculovirus, the baculovirus flanking sequences necessary for

proper cross-over during recombination (the flanking sequences comprise about 200-300 base pairs adjacent to the promoter sequence) and a bacterial origin of replication which permits the construct to replicate in bacteria. The vector is constructed so that (i) the DNA segment is placed adjacent (or operably-linked or "downstream" or "under the control of") to the polyhedron gene promoter and (ii) the promoter/dehydrodiconiferyl alcohol benzylic ether reductase combination is flanked on both sides by 200-300 base pairs of baculovirus DNA (the flanking sequences).

To produce a dehydrodiconiferyl alcohol benzylic ether reductase DNA construct, a cDNA clone encoding a full length dehydrodiconiferyl alcohol benzylic ether reductase is obtained using methods such as those described herein. The DNA construct is contacted in a host cell with baculovirus DNA of an appropriate baculovirus (that is, of the same species of baculovirus as the promoter encoded in the construct) under conditions such that recombination is effected. The resulting recombinant baculoviruses encode the full dehydrodiconiferyl alcohol benzylic ether reductase. For example, an insect host cell can be cotransfected or transfected separately with the DNA construct and a functional baculovirus. Resulting recombinant baculoviruses can then be isolated and used to infect cells to effect production of dehydrodiconiferyl alcohol benzylic ether reductase. Host insect cells include, for example, *Spodoptera frugiperda* cells. Insect host cells infected with a recombinant baculovirus of the present invention are then cultured under conditions allowing expression of the baculovirus-encoded dehydrodiconiferyl alcohol benzylic ether reductase. Recombinant protein thus produced is then extracted from the cells using methods known in the art.

Other eukaryotic microbes such as yeasts may also be used to practice this invention. The baker's yeast *Saccharomyces cerevisiae*, is a commonly used yeast, although several other strains are available. The plasmid YRp7 (Stinchcomb et al., *Nature* **282:**39 (1979); Kingsman et al., *Gene* **7:**141 (1979); Tschemper et al., *Gene* **10:**157 (1980)) is commonly used as an expression vector in *Saccharomyces*. This plasmid contains the trp1 gene that provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, such as strains ATCC No. 44,076 and PEP4-1 (Jones, *Genetics* **85:**12 (1977)). The presence of the trp1 lesion as a characteristic of the yeast host cell genome then provides an effective environment for detecting transformation by growth in the absence of tryptophan. Yeast host cells are generally transformed using the polyethylene glycol method, as described by Hinnen (*Proc. Natl. Acad. Sci. USA* **75:**1929 (1978)). Additional yeast

transformation protocols are set forth in Gietz et al., *N.A.R.* **20(17):**1425 (1992); Reeves et al., *FEMS* **99:**193-197 (1992).

Suitable promoting sequences in yeast vectors include the promoters for 3-phosphoglycerate kinase (Hitzeman et al., *J. Biol. Chem.* **255:**2073 (1980)) or other glycolytic enzymes (Hess et al., *J. Adv. Enzyme Reg.* **7:**149 (1968); Holland et al., *Biochemistry* **17:**4900 (1978)), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triose-phosphate isomerase, phosphoglucose isomerase, and glucokinase. In the construction of suitable expression plasmids, the termination sequences associated with these genes are also ligated into the expression vector 3' of the sequence desired to be expressed to provide polyadenylation of the mRNA and termination. Other promoters that have the additional advantage of transcription controlled by growth conditions are the promoter region for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and the aforementioned glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Any plasmid vector containing yeast-compatible promoter, origin of replication and termination sequences is suitable.

Cell cultures derived from multicellular organisms, such as plants, may be used as hosts to practice this invention. Transgenic plants can be obtained, for example, by transferring plasmids that encode dehydrodiconiferyl alcohol benzylic ether reductase, and a selectable marker gene, e.g., the kan gene encoding resistance to kanamycin, into *Agrobacterium tumifaciens* containing a helper Ti plasmid as described in Hoeckema et al., *Nature* **303:**179-181 (1983) and culturing the *Agrobacterium* cells with leaf slices of the plant to be transformed as described by An et al., *Plant Physiology* **81:**301-305 (1986). Transformation of cultured plant host cells is normally accomplished through *Agrobacterium tumifaciens*, as described above. Cultures of mammalian host cells and other host cells that do not have rigid cell membrane barriers are usually transformed using the calcium phosphate method as originally described by Graham and Van der Eb (*Virology* **52:**546 (1978)) and modified as described in Sections 16.32-16.37 of Sambrook et al., *supra*. However, other methods for introducing DNA into cells such as Polybrene (Kawai and Nishizawa, *Mol. Cell. Biol.* **4:**1172 (1984)), protoplast fusion (Schaffner, *Proc. Natl. Acad. Sci. USA* 77:2163 (1980)), electroporation (Neumann et al., *EMBO J.* **1:**841 (1982)), and direct microinjection into nuclei (Capecchi, *Cell* **22:**479 (1980)) may

also be used. Additionally, animal transformation strategies are reviewed in Monastersky G.M. and Robl, J.M., *Strategies in Transgenic Animal Science*, ASM Press, Washington, D.C. (1995). Transformed plant calli may be selected through the selectable marker by growing the cells on a medium containing, e.g., kanamycin, and appropriate amounts of phytohormone such as naphthalene acetic acid and benzyladenine for callus and shoot induction. The plant cells may then be regenerated and the resulting plants transferred to soil using techniques well known to those skilled in the art.

In addition, a nucleic acid sequence encoding dehydrodiconiferyl alcohol benzylic ether reductase can be incorporated into the plant along with a necessary promoter which is inducible. In the practice of this embodiment of the invention, a promoter that only responds to a specific external or internal stimulus is fused to the target cDNA. Thus, the gene will not be transcribed except in response to the specific stimulus. As long as the gene is not being transcribed, its gene product is not produced.

An illustrative example of a responsive promoter system that can be used in the practice of this invention is the glutathione-S-transferase (GST) system in maize. GSTs are a family of enzymes that can detoxify a number of hydrophobic electrophilic compounds that often are used as pre-emergent herbicides (Weigand et al., *Plant Molecular Biology* 7:235-243 (1986)). Studies have shown that the GSTs are directly involved in causing this enhanced herbicide tolerance. This action is primarily mediated through a specific 1.1 kb mRNA transcription product. In short, maize has a naturally occurring quiescent gene already present that can respond to external stimuli and that can be induced to produce a gene product. This gene has previously been identified and cloned. Thus, in one embodiment of this invention, the promoter is removed from the GST responsive gene and attached to a dehydrodiconiferyl alcohol benzylic ether reductase gene that previously has had its native promoter removed. This engineered gene is the combination of a promoter that responds to an external chemical stimulus and a gene responsible for successful production of dehydrodiconiferyl alcohol benzylic ether reductase protein.

In addition to the methods described above, several methods are known in the art for transferring cloned DNA into a wide variety of plant species, including gymnosperms, angiosperms, monocots and dicots (see, e.g., Glick and Thompson, eds., *Methods in Plant Molecular Biology*, CRC Press, Boca Raton, Florida (1993)). Representative examples include electroporation-facilitated DNA uptake by

protoplasts (Rhodes et al., *Science* **240(4849):**204-207 (1988)); treatment of protoplasts with polyethylene glycol (Lyznik et al., *Plant Molecular Biology* **13:**151-161 (1989)); and bombardment of cells with DNA laden microprojectiles (Klein et al., *Plant Physiol.* **91:**440-444 (1989) and Boynton et al., *Science* **240(4858):**1534-1538 (1988)). Numerous methods now exist, for example, for the transformation of cereal crops (see, e.g., McKinnon, G.E. and Henry, R.J., *J. Cereal Science* **22(3):**203-210 (1995); Mendel, R.R. and Teeri, T.H., *Plant and Microbial Biotechnology Research Series*, 3:81-98, Cambridge University Press (1995); McElroy, D. and Brettell, R.I.S., *Trends in Biotechnolog*, **12(2):**62-68 (1994); Christou et al., *Trends in Biotechnology* **10(7):**239-246 (1992); Christou, P. and Ford, T.L., *Annals of Botan*, **75(5):**449-454 (1995); Park et al., *Plant Molecular Biolog*, **32(6):**1135-1148 (1996); Altpeter et al., *Plant Cell Report*, **16:**12-17 (1996)). Additionally, plant transformation strategies and techniques are reviewed in Birch, R.G., *Ann Rev Plant Phys Plant Mol Biol* **48:**297 (1997); Forester et al., *Exp. Agric.* **33:**15-33 (1997). Minor variations make these technologies applicable to a broad range of plant species.

Each of these techniques has advantages and disadvantages. In each of the techniques, DNA from a plasmid is genetically engineered such that it contains not only the gene of interest, but also selectable and screenable marker genes. A selectable marker gene is used to select only those cells that have integrated copies of the plasmid (the construction is such that the gene of interest and the selectable and screenable genes are transferred as a unit). The screenable gene provides another check for the successful culturing of only those cells carrying the genes of interest. A commonly used selectable marker gene is neomycin phosphotransferase II (NPT II). This gene conveys resistance to kanamycin, a compound that can be added directly to the growth media on which the cells grow. Plant cells are normally susceptible to kanamycin and, as a result, die. The presence of the NPT II gene overcomes the effects of the kanamycin and each cell with this gene remains viable. Another selectable marker gene which can be employed in the practice of this invention is the gene which confers resistance to the herbicide glufosinate (Basta). A screenable gene commonly used is the β-glucuronidase gene (GUS). The presence of this gene is characterized using a histochemical reaction in which a sample of putatively transformed cells is treated with a GUS assay solution. After an appropriate incubation, the cells containing the GUS gene turn blue. Preferably, the plasmid will contain both selectable and screenable marker genes.

The plasmid containing one or more of these genes is introduced into either plant protoplasts or callus cells by any of the previously mentioned techniques. If the marker gene is a selectable gene, only those cells that have incorporated the DNA package survive under selection with the appropriate phytotoxic agent. Once the appropriate cells are identified and propagated, plants are regenerated. Progeny from the transformed plants must be tested to insure that the DNA package has been successfully integrated into the plant genome.

Mammalian host cells may also be used in the practice of the invention. Examples of suitable mammalian cell lines include monkey kidney CVI line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line 293S (Graham et al., *J. Gen. Virol.* **36**:59 (1977)); baby hamster kidney cells (BHK, ATCC CCL 10); Chinese hamster ovary cells (Urlab and Chasin, *Proc. Natl. Acad. Sci USA* 77:4216 (1980)); mouse sertoli cells (TM4, Mather, *Biol. Reprod.* **23**:243 (1980)); monkey kidney cells (CVI-76, ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor cells (MMT 060562, ATCC CCL 51); rat hepatoma cells (HTC, MI.54, Baumann et al., *J. Cell Biol.* **85**:1 (1980)); and TRI cells (Mather et al., *Annals N.Y. Acad. Sci.* **383**:44 (1982)). Expression vectors for these cells ordinarily include (if necessary) DNA sequences for an origin of replication, a promoter located in front of the gene to be expressed, a ribosome binding site, an RNA splice site, a polyadenylation site, and a transcription terminator site.

Promoters used in mammalian expression vectors are often of viral origin. These viral promoters are commonly derived from polyoma virus, Adenovirus 2, and most frequently Simian Virus 40 (SV40). The SV40 virus contains two promoters that are termed the early and late promoters. These promoters are particularly useful because they are both easily obtained from the virus as one DNA fragment that also contains the viral origin of replication (Fiers et al., *Nature* **273**:113 (1978)). Smaller or larger SV40 DNA fragments may also be used, provided they contain the approximately 250-bp sequence extending from the HindIII site toward the BglI site located in the viral origin of replication.

Alternatively, promoters that are naturally associated with the foreign gene (homologous promoters) may be used provided that they are compatible with the host cell line selected for transformation.

An origin of replication may be obtained from an exogenous source, such as SV40 or other virus (e.g., Polyoma, Adeno, VSV, BPV) and inserted into the cloning vector. Alternatively, the origin of replication may be provided by the host cell chromosomal replication mechanism. If the vector containing the foreign gene is integrated into the host cell chromosome, the latter is often sufficient.

The use of a secondary DNA coding sequence can enhance production levels of dehydrodiconiferyl alcohol benzylic ether reductase protein in transformed cell lines. The secondary coding sequence typically comprises the enzyme dihydrofolate reductase (DHFR). The wild-type form of DHFR is normally inhibited by the chemical methotrexate (MTX). The level of DHFR expression in a cell will vary depending on the amount of MTX added to the cultured host cells. An additional feature of DHFR that makes it particularly useful as a secondary sequence is that it can be used as a selection marker to identify transformed cells. Two forms of DHFR are available for use as secondary sequences, wild-type DHFR and MTX-resistant DHFR. The type of DHFR used in a particular host cell depends on whether the host cell is DHFR deficient (such that it either produces very low levels of DHFR endogenously, or it does not produce functional DHFR at all). DHFR-deficient cell lines such as the CHO cell line described by Urlaub and Chasin, *supra*, are transformed with wild-type DHFR coding sequences. After transformation, these DHFR-deficient cell lines express functional DHFR and are capable of growing in a culture medium lacking the nutrients hypoxanthine, glycine and thymidine. Nontransformed cells will not survive in this medium.

The MTX-resistant form of DHFR can be used as a means of selecting for transformed host cells in those host cells that endogenously produce normal amounts of functional DHFR that is MTX sensitive. The CHO-Kl cell line (ATCC No. CL 61) possesses these characteristics, and is thus a useful cell line for this purpose. The addition of MTX to the cell culture medium will permit only those cells transformed with the DNA encoding the MTX-resistant DHFR to grow. The nontransformed cells will be unable to survive in this medium.

Prokaryotes may also be used as host cells for the initial cloning steps of this invention. They are particularly useful for rapid production of large amounts of DNA, for production of single-stranded DNA templates used for site-directed mutagenesis,

for screening many mutants simultaneously, and for DNA sequencing of the mutants generated. Suitable prokaryotic host cells include *E. coli* K12 strain 294 (ATCC No. 31,446), *E. coli* strain W3110 (ATCC No. 27,325) *E. coli* X1776 (ATCC No. 31,537), and *E. coli* B; however many other strains of *E. coli*, such as HB101, JM101, NM522, NM538, NM539, and many other species and genera of prokaryotes including bacilli such as *Bacillus subtilis*, other enterobacteriaceae such as *Salmonella typhimurium* or *Serratia marcesans*, and various *Pseudomonas* species may all be used as hosts. Prokaryotic host cells or other host cells with rigid cell walls are preferably transformed using the calcium chloride method as described in section 1.82 of Sambrook et al., *supra*. Alternatively, electroporation may be used for transformation of these cells. Prokaryote transformation techniques are set forth in Dower, W. J., in Genetic Engineering, Principles and Methods, **12**:275-296, Plenum Publishing Corp. (1990); Hanahan et al., *Meth. Enzymol.*, **204**:63 (1991).

As a representative example, cDNA sequences encoding dehydrodiconiferyl alcohol benzylic ether reductase may be transferred to the $(His)_6$•Tag pET vector commercially available (from Novagen) for overexpression in *E. coli* as heterologous host. This pET expression plasmid has several advantages in high level heterologous expression systems. The desired cDNA insert is ligated in frame to plasmid vector sequences encoding six histidines followed by a highly specific protease recognition site (thrombin) that are joined to the amino terminus codon of the target protein. The histidine "block" of the expressed fusion protein promotes very tight binding to immobilized metal ions and permits rapid purification of the recombinant protein by immobilized metal ion affinity chromatography. The histidine leader sequence is then cleaved at the specific proteolysis site by treatment of the purified protein with thrombin, and the dehydrodiconiferyl alcohol benzylic ether reductase protein eluted. This overexpression-purification system has high capacity, excellent resolving power and is fast, and the chance of a contaminating *E. coli* protein exhibiting similar binding behavior (before and after thrombin proteolysis) is extremely small.

As will be apparent to those skilled in the art, any plasmid vectors containing replicon and control sequences that are derived from species compatible with the host cell may also be used in the practice of the invention. The vector usually has a replication site, marker genes that provide phenotypic selection in transformed cells, one or more promoters, and a polylinker region containing several restriction sites for insertion of foreign DNA. Plasmids typically used for transformation of *E. coli* include pBR322, pUC18, pUC19, pUCI18, pUC119, and Bluescript M13, all of

which are described in Sections 1.12-1.20 of Sambrook et al., *supra*. However, many other suitable vectors are available as well. These vectors contain genes coding for ampicillin and/or tetracycline resistance which enables cells transformed with these vectors to grow in the presence of these antibiotics.

The promoters most commonly used in prokaryotic vectors include the β-lactamase (penicillinase) and lactose promoter systems (Chang et al. *Nature* **375**:615 (1978); Itakura et al., *Science* **198**:1056 (1977); Goeddel et al., *Nature* **281**:544 (1979)) and a tryptophan (trp) promoter system (Goeddel et al., *Nucl. Acids Res.* **8**:4057 (1980); EPO Appl. Publ. No. 36,776), and the alkaline phosphatase systems. While these are the most commonly used, other microbial promoters have been utilized, and details concerning their nucleotide sequences have been published, enabling a skilled worker to ligate them functionally into plasmid vectors (see Siebenlist et al., *Cell* **20**:269 (1980)).

Many eukaryotic proteins normally secreted from the cell contain an endogenous secretion signal sequence as part of the amino acid sequence. Thus, proteins normally found in the cytoplasm can be targeted for secretion by linking a signal sequence to the protein. This is readily accomplished by ligating DNA encoding a signal sequence to the 5' end of the DNA encoding the protein and then expressing this fusion protein in an appropriate host cell. The DNA encoding the signal sequence may be obtained as a restriction fragment from any gene encoding a protein with a signal sequence. Thus, prokaryotic, yeast, and eukaryotic signal sequences may be used herein, depending on the type of host cell utilized to practice the invention. The DNA and amino acid sequence encoding the signal sequence portion of several eukaryotic genes including, for example, human growth hormone, proinsulin, and proalbumin are known (see Stryer, *Biochemistry* W.H. Freeman and Company, New York, NY, p. 769 (1988)), and can be used as signal sequences in appropriate eukaryotic host cells. Yeast signal sequences, as for example acid phosphatase (Arima et al., *Nucleic Acids Res.* **11**:1657 (1983)), alpha-factor, alkaline phosphatase and invertase may be used to direct secretion from yeast host cells. Prokaryotic signal sequences from genes encoding, for example, LamB or OmpF (Wong et al., *Gene* **68**:193 (1988)), MalE, PhoA, or beta-lactamase, as well as other genes, may be used to target proteins from prokaryotic cells into the culture medium.

Trafficking sequences from plants, animals and microbes can be employed in the practice of the invention to direct the gene product to the cytoplasm, endoplasmic reticulum, mitochondria or other cellular components, or to target the protein for

export to the medium. These considerations apply to the overexpression of dehydrodiconiferyl alcohol benzylic ether reductase, and to direction of expression within cells or intact organisms to permit gene product function in any desired location.

The construction of suitable vectors containing DNA encoding replication sequences, regulatory sequences, phenotypic selection genes and the dehydrodiconiferyl alcohol benzylic ether reductase DNA of interest are prepared using standard recombinant DNA procedures. Isolated plasmids and DNA fragments are cleaved, tailored, and ligated together in a specific order to generate the desired vectors, as is well known in the art (see, for example, Sambrook et al., *supra*).

As discussed above, dehydrodiconiferyl alcohol benzylic ether reductase variants, are preferably produced by means of mutation(s) that are generated using the method of site-specific mutagenesis. This method requires the synthesis and use of specific oligonucleotides that encode both the sequence of the desired mutation and a sufficient number of adjacent nucleotides to allow the oligonucleotide to stably hybridize to the DNA template.

The foregoing may be more fully understood in connection with the following representative examples, in which "Plasmids" are designated by a lower case p followed by an alphanumeric designation. The starting plasmids used in this invention are either commercially available, publicly available on an unrestricted basis, or can be constructed from such available plasmids using published procedures. In addition, other equivalent plasmids are known in the art and will be apparent to the ordinary artisan.

"Digestion", "cutting" or "cleaving" of DNA refers to catalytic cleavage of the DNA with an enzyme that acts only at particular locations in the DNA. These enzymes are called restriction endonucleases, and the site along the DNA sequence where each enzyme cleaves is called a restriction site. The restriction enzymes used in this invention are commercially available and are used according to the instructions supplied by the manufacturers. (See also Sections 1.60-1.61 and Sections 3.38-3.39 of Sambrook et al., *supra*.).

"Recovery" or "isolation" of a given fragment of DNA from a restriction digest means separation of the resulting DNA fragment on a polyacrylamide or an agarose gel by electrophoresis, identification of the fragment of interest by comparison of its mobility versus that of marker DNA fragments of known molecular weight, removal of the gel section containing the desired fragment, and separation of

the gel from DNA. This procedure is known generally. For example, see Lawn et al. (*Nucleic Acids Res.* 9:6103-6114 (1982)), and Goeddel et al. (*Nucleic Acids Res., supra*).

EXAMPLE 1

Cloning a Dehydrodiconiferyl alcohol benzylic ether reductase protein from *Pinus taeda*

Unless otherwise indicated, the following materials, methods and instrumentation were used in Example 1 and all succeeding examples.

Plant Materials — *P. taeda* cell suspension cultures were maintained as described previously (van Heerden, P.S., Towers, G.H.N. & Lewis, N.G. *J. Biol. Chem.* **271**, 12350-12355 (1996)), in media containing 2,4-dichlorophenoxyacetic acid (2,4-D). Cells were harvested by filtration seven days after transfer to fresh media, frozen in liquid nitrogen and stored at –80 °C.

General Methods — All molecular biological techniques, unless expressly described below were performed according to standard methods (Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual, 3 volumes, 3rd Ed.* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1994); Ausubel, F.M., *et al. Current Protocols in Molecular Biology, 2 volumes* (Greene Publishing Associates and Wiley-Interscience, John Wiley & Sons, NY, 1991)).

Materials — All solvents and chemicals used were reagent or HPLC grade. *Taq* thermostable DNA polymerase was obtained from Promega. Competent NovaBlue cells were purchased from Novagen and radiolabeled nucleotide ([α^{-32}P]dCTP) was from DuPont NEN. The pCRII TA cloning kit was from Invitrogen. Restriction endonuclease NdeI was from New England Biolabs.

Oligonucleotide primers for polymerase chain reaction (PCR) and sequencing were synthesized by Gibco BRL Life Technologies. GENECLEAN II® kits (BIO 101 Inc.) were used for purification of PCR fragments, with the gel purified DNA concentrations determined by comparison to a low DNA mass ladder (Gibco BRL) in 1.5% agarose gels.

Instrumentation — UV (including RNA and DNA determinations at OD_{260}) spectra were recorded on a Lambda 6 UV/VIS spectrophotometer. A Temptronic II thermocycler (Thermolyne) was used for all PCR amplifications. Purification of plasmid DNA for sequencing employed a QIAwell Plus plasmid

purification system (Qiagen) followed by PEG precipitation or Wizard® Plus SV Minipreps DNA Purification System (Promega), with DNA sequences determined using an Applied Biosystems Model 373A automated sequencer. All high performance liquid chromatographic (HPLC) separations were performed on either a Millenium (Waters, Inc.) or Alliance (Waters, Inc) instrument, with eluent monitoring at 280 nm.

Pinus taeda cDNA Library Synthesis — Total RNA (100 µg/g fresh weight) was obtained using the method of Dong and Dunsten (Dong, J.Z. & Dunstan, D.I. *Plant Cell Reports* **15** (1996)) from frozen loblolly pine (*Pinus taeda*) cells grown as described above. A *P. taeda* cDNA library was constructed using 5 µg of purified poly(A)$^+$ mRNA (Oligotex-dT™ Suspension, Qiagen) with the ZAP-cDNA® synthesis kit, the Uni ZAP™ XR vector, and the Gigapack® II Gold packaging extract (Stratagene), with a titer of 1 X 10^6 pfu for the primary library. The amplified library (1 X 10^9 pfu /ml; 120 ml total) was used for screening (Dinkova-Kostova, A.T., *et al.* *J. Biol. Chem.* **271**, 29473-29482 (1996)).

DNA Probe Synthesis — The 5′-end of a previously isolated pinoresinol-lariciresinol reductase cDNA (PLR-Fi1) (Dinkova-Kostova, A.T., *et al. J. Biol. Chem.* **271**:29473-29482 (1996)), having the nucleic acid sequence set forth in (SEQ ID NO:7), was used as a probe to screen the *P. taeda* cDNA library for similar/homologous enzymes. The probe was constructed as follows: 10 ng of purified pBSPLR-Fi1 plasmid (PLR-Fi1 contained in the cloning plasmid pBluescript SK[-]) was used as the template in five 100 µl PCR reactions (10 mM Tris-HCl [pH 9.0], 50 mM KCl, 0.1% Triton X-100, 2.5 mM $MgCl_2$, 0.2 mM each dNTP and 2.5 units *Taq* DNA polymerase) with primers PLRNT1 (AT(A/T/C) AT(A/T/C) GGI GGI ACI GGI TA) (SEQ ID NO:8) (100 pmol) and PLRI5R (TC(T/C) TCI A(A/G)I GTI AC(T/C) TTI CC) (SEQ ID NO:9) (100 pmol). The PCR amplifications were carried out in a thermocycler as follows: 35 cycles of 1 min at 94°C, 2 min at 50°C and 3 min at 72°C; with 5 min at 72°C and an indefinite hold at 4°C after the final cycle. The 5 reactions were concentrated (Microcon 30, Amicon Inc.) and washed with TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA; 2 x 200 µl), with the PCR product subsequently recovered in TE buffer (2 x 50 µl). The PLR-Fi1 5′-end reaction product (~400 bp band) was resolved in a preparative 1.5% agarose gel and purified from the agarose using the GENECLEAN II® kit (BIO 101 Inc.). Gel purified PLR-Fi1 5′-end fragment (SEQ ID NO:7) (50 ng) was used with Pharmacia's T7QuickPrime® kit and [α-^{32}P]dCTP, according to kit instructions, to

produce a radiolabeled probe (in 0.1 ml), which was purified over BioSpin 6 columns (Bio-Rad) and added to carrier DNA (0.9 ml of 0.5 mg/ml sheared salmon sperm DNA [Sigma]).

Library Screening — 600,000 pfu of *P. taeda* amplified cDNA library were plated for primary screening, according to Stratagene's instructions. Plaques were blotted onto Magna Nylon membrane circles (Micron Separations Inc.), which were then allowed to air dry. The membranes were placed between two layers of Whatman® 3MM Chr paper. cDNA library phage DNA was fixed to the membranes and denatured in one step by autoclaving for 2 min at 100°C with fast exhaust. The membranes were washed for 30 min at 37°C in 6X standard saline citrate (SSC) and 0.1% SDS and prehybridized for 5 h with gentle shaking at 45 °C in preheated 6X SSC, 0.5% SDS and 5X Denhardt's reagent (hybridization solution, 300 ml) in a crystallization dish (190 x 75 mm). The [^{32}P]radiolabeled probe (SEQ ID NO:7) was denatured (boiling, 10 min), quickly cooled (ice, 15 min) and added to a preheated fresh hybridization solution (60 ml, 45 °C) in a crystallization dish (150 x 75 mm). The prehybridized membranes were next added to this dish, which was then covered with plastic wrap. Hybridization was performed for 18 h at 45°C with gentle shaking. The membranes were washed in 4X SSC and 0.5% SDS for 5 min at room temperature, transferred to 2X SSC and 0.5% SDS (at room temperature) and incubated at 45 °C for 20 min with gentle shaking, wrapped with plastic wrap to prevent drying and finally exposed to Kodak X-OMAT AR film for 24 h at -80°C with intensifying screens. Twenty positive plaques were purified through two more rounds of screening with hybridization conditions as above.

In vivo Excision and Sequencing of Putative Reductase Protein cDNA-containing Phagemids — Purified cDNA clones were rescued from the phage following Stratagene's *in vivo* excision protocol. Both strands of several different cDNAs that coded for genes homologous to PLR-Fi1 were completely sequenced using overlapping sequencing primers. Two cDNAs were identified, but which only differed in their 5′- and 3′-untranslated regions at the site of insertion into the pBluescript SK[-] cloning plasmid. The differences between these two cDNAs were cloning artifacts, therefore, only a single gene homologous to the *F. intermedia* pinoresinol-lariciresinol reductase was cloned from this *P. taeda* cell suspension culture cDNA library. The nucleotide sequence of this reductase gene is shown in (SEQ ID NO:1) which sets forth the sequence that is common to both of the two, foregoing, reductase cDNAs.

Sequence Analysis — DNA and amino acid sequence analyses were performed using the Unix-based GCG Wisconsin Package (*Genetics Computer Group*, 575 Science Drive, Madison, Wisconsin, USA, 1994); Rice, P. (The Sanger Centre, Hinxton Hall, Cambridge, England (1996)) and the ExPASy World Wide Web molecular biology server (Geneva University Hospital and University of Geneva, Geneva, Switzerland).

EXAMPLE 2

Expression of Dehydrodiconiferyl alcohol benzylic ether reductase protein in E. coli as a fusion protein

Expression of Dehydrodiconiferyl alcohol benzylic ether reductase protein in Escherichia coli as a Fusion Protein —The open reading frame of the putative reductase from *P. taeda* was in frame with the β-galactosidase gene α-complementation particle in pBluescript. Thus, its purified plasmid DNA was transformed into NovaBlue cells according to Novagen's instructions. Transformed cells (5 ml cultures) were grown at 37°C with shaking (225 rpm) to mid log phase (OD_{600}=0.5) in LB medium supplemented with 12.5 μg ml tetracycline and 50 μg ml ampicillin. IPTG (isopropyl β-D-thioglucopyranoside) was then added to a final concentration of 10 mM, and the cells were allowed to grow for 2 h. Cells were collected by centrifugation and resuspended in 500 μl (per 5 ml culture tube) buffer (20 mM Tris-HCl, pH 8.0, 5 mM dithiothreitol). Lysozyme (5 μl of 0.1 mg ml^{-1}, Research Organics, Inc.) was next added and following incubation for 10 min, the cells were lysed by sonication (3 x 15 seconds). After centrifugation at 14,000 x g at 4°C for 10 min, the supernatant was removed and assayed for both pinoresinol-lariciresinol reductase and dehydrodiconiferyl alcohol reductase activities (210 μl supernatant per assay) as described herein and in Example 3. No pinoresinol-lariciresinol reductase activity was observed in *E. coli* extracts expressing the protein as a fusion with the β-galactosidase gene α-complementation particle, even when the assays were allowed to incubate for 24 h, although an identical system has been used to produce catalytically active (+)-pinoresinol/(+)-lariciresinol reductase from *F. intermedia* (Dinkova-Kostova, A.T., *et al. J. Biol. Chem.* **271**:29473-29482 (1996)).

Radiochemical Assays for Pinoresinol-Lariciresinol Reductase Activity — Pinoresinol reductase activity was evaluated by monitoring formation of [^{3}H]lariciresinol and [^{3}H]secoisolariciresinol. Each assay for pinoresinol reductase activity consisted of 20 mM Bis-Tris Propane, pH 7.0, 0.4 mM (±)-pinoresinols

(added in 20 µl MeOH) and the enzyme preparation (i.e., total protein extract from *E. coli*, 210 µl). The enzymatic reaction was initiated by addition of 0.8 mM [4*R*-^{3}H]NADPH (6.79 MBq/mmol). After 3 hour incubation at 30°C with shaking, the assay mixture was extracted with EtOAc (500 µl) containing (±)-lariciresinols (20 µg) and (±)-secoisolariciresinols (20 µg) as radiochemical carriers. After centrifugation (17,000 × g, 5 min), the EtOAc solubles were removed and the extraction procedure was continued with 500 µl EtOAc. For each assay, the EtOAc solubles were combined, with an aliquot (100 µl) removed for determination of its radioactivity using liquid scintillation counting. The remainder of the combined EtOAc solubles was evaporated to dryness *in vacuo*, reconstituted in MeOH/H_2O (3:7, 100 µl) and subjected to reversed phase HPLC. Lariciresinol reductase activity was evaluated by monitoring formation of [^{3}H]secoisolariciresinol. Assays were carried out exactly as described above, except that 0.4 mM (±)-lariciresinols were used as substrates, with (±)-secoisolariciresinols (20 µg) added as radiochemical carriers. HPLC was carried out as previously described (Dinkova-Kostova, A.T., *et al. J. Biol. Chem.* 271:29473-29482 (1996)) to separate the lignan substrates and products. Briefly, reversed-phase column chromatography employed a Nova-pak C_{18} column (3.5 mm × 150 mm, Waters) with an isocratic solvent system consisting of MeOH:3 % acetic acid (in H_2O) (30:70) at a flow rate of 0.5 ml min^{-1}. 0.5 ml fractions were collected for scintillation counting to determine the level of incorporation of ^{3}H into the assay products.

Non-radioactive Assays for Pinoresinol-Lariciresinol Reductase Activity — Pinoresinol reductase activity was further evaluated in assays performed as above with the following exceptions: total volume was 150 µl; 4 mM NADPH (not radioactive) was used; 2 mM (±)-pinoresinols or (±)-lariciresinols were used as substrates; no (±)-lariciresinols or (±)-secoisolariciresinols were added as radiochemical carriers; the reactions were stopped by boiling for 3 min; 50 µl MeOH was then added to bring the concentration to 30 % and the assay mixtures were centrifuged (17,000 ×g, 3 min); 150 µl of the resulting mixture was directly injected on the HPLC to determine whether any lariciresinol and secoisolariciresinol was formed.

It was found that when assays were allowed to continue for up to 24 hours (i.e. under conditions that would have depleted all available substrate with the *F. intermedia* reductase [PLR-Fi1] (Dinkova-Kostova, A.T., *et al. J. Biol. Chem.* 271:29473-29482 (1996)), neither pinoresinol nor lariciresinol were reduced to give either lariciresinol or secoisolariciresinol, respectively.

EXAMPLE 3

Expression of Native Dehydrodiconiferyl alcohol benzylic ether reductase protein (SEQ ID NO:2) in *E. coli*

Transfer of the Putative Reductase into pSBETa — The over-expression plasmid, pSBETa, contains the pET3a expression cassette, with two restriction sites available for sub-cloning (*Nde*I for native expression and *BamH*I for expression with a small fusion), as well as the *argU* gene (for production of the rare AGA-Arg $tRNA_{arg4}$) and kanamycin resistance (for high plasmid stability in liquid cultures). Since the putative reductase from *P. taeda* had no internal *Nde*I sites, two primers were designed to introduce *Nde*I sites at the start methionine (primer PT-ATG-NdeI: TTC AGG GCC CAT ATG GGA AGC AGG AGC AGG ATA CTC) (SEQ ID NO:10) and in the 3′-end untranslated region (primer PT-Rev-NdeI: TGT CGA ATA CAT ATG AAA GGC GAT AAC CAA CAA TTT) (SEQ ID NO:11). These two primers (SEQ ID NO:10) and (SEQ ID NO:11) (5 pmol each) were used in five PCR reactions with 10 ng of the cDNA (SEQ ID NO:1) encoding the *P. taeda* putative reductase (SEQ ID NO:2) as described above and sub-cloned into the pCRII plasmid according to Invitrogen's instructions. The resulting pCRII containing the putative reductase was purified and digested with *Nde*I, with the resulting ~1 kb fragment gel purified and ligated into pSBETa previously digested with *Nde*I and transformed into competent NovaBlue cells. The resulting pSBET construct containing the putative reductase was purified and the expression region, containing the desired cDNA, was sequenced completely on both strands to verify that no mutations had been introduced during PCR.

General Procedures for Enzyme *Purification* — Protein purification procedures were carried out at 4 °C, with chromatographic eluents monitored at 280 nm. Protein concentrations were determined with BioRad's protein determination kit. Polyacrylamide gel electrophoresis used gradient (4-15%, linear gradient, BioRad) gels under denaturing and reducing conditions in the Laemmli buffer system, followed by visualization of the proteins by silver staining.

Over-Expression of the Native Enzyme (SEQ ID NO:2) in E. coli — The resulting pSBETa plasmid containing the putative reductase from *P. taeda* (SEQ ID NO:2) was transformed into B834(DE3) *E. coli* cells for expression. High level expression of the putative reductase (SEQ ID NO:2) was achieved by inoculating 1 liter of LB broth, supplemented with 50 μg·ml kanamycin, with 2-4 ml of an overnight grown 10 ml culture in the same medium. The cells were then allowed to grow at 37

°C with shaking at 250 rpm until a density of OD_{600}=0.65 was reached, at which point the growth conditions were changed to 20 °C with shaking at 265 rpm. Production of the reductase (SEQ ID NO:2) was induced by addition of IPTG to 1 mM final concentration, once the cells had reached a temperature of ~22 °C (about 30 min after the incubator temperature was turned down). The cells were then allowed to grow for 21 h before harvest by centrifugation for 25 min at 3000 ×g in 4 × 250 ml centrifuge bottles, after which the pellets were frozen for at least 2 h at –80 °C, to aid in cell lysis.

Crude Protein Preparation — The four pellets of *E. coli* cells obtained in the previous step were thawed at room temperature and resuspended in 2× 10 ml each of Buffer A (20 mM Tris·HCl, pH 8.0; 2 mM ethylenediamine tetraacetic acid [EDTA]; 1 mM phenylmethylsulfonyl fluoride [PMSF]; (Paré, P.W., Wang, H.-B., Davin, L.B. & Lewis, N.G. *Tetrahedron Lett.* **35**:4731-4734 (1994)) 5 mM dithiothreitol [DTT]), then combined and sonicated for 3 × 30 s. The sonicates were centrifuged for 30 min at 20,000 ×g to pellet cellular debris and filtered through a 0.2 µm syringe filter. The resulting filtrate (290 ml) was subjected to ammonium sulfate precipitation, with the 40-70 % ammonium sulfate saturation cut, after centrifugation for 30 min at 20,000 ×g, desalted back into Buffer A over PD-10 desalting columns (Pharmacia).

Affinity Column (Affi-Blue gel) Purification — A 1.6 cm diameter, 11.5 cm long (23 ml bed volume) Affi-Gel blue gel (BioRad) column was pre-equilibrated in Buffer A. The desalted 40-70% cut (15 ml, 513 mg) was then applied to the column. After washing with 300 ml Buffer A (1 $ml \cdot min^{-1}$), the reductase (SEQ ID NO:2) was eluted by running a linear gradient from 0 – 100 % Buffer B (Buffer A + 5 M NaCl) in 500 mL, holding at 100 % Buffer B for 60 mL and returning from 100 % Buffer B to 100 % Buffer A in 30 ml. 4 ml fractions were collected, with the putative reductase (SEQ ID NO:2) being eluted in fractions 33-43. These fractions were then pooled, concentrated to ~ 5mg/ml in a Centricon 10 microconcentrator (Amicon, Inc.), desalted over PD10 columns and assayed for activity.

Anion Exchange Chromatography — The resulting enzyme solution (118 mg) was next applied to a POROS 20 QE Perfusion anion exchange column, pre-equilibrated in Buffer C (50 mM Bis-Tris Propane, pH 6.8, 5 mM DTT), and eluted in the load (i.e., did not bind to the column), whereas most of the contaminating *E. coli* proteins remained bound to the anionic exchange column under these conditions.

Cationic Exchange Chromatography — The resulting enzyme solution (59 mg) was next applied to a POROS 20 SP Perfusion cationic exchange column, pre-

equilibrated in Buffer C (50 mM Bis-Tris Propane, pH 6.8, 5 mM DTT), and eluted in the load (i.e., did not bind to the column), whereas most of the rest of the contaminating *E. coli* proteins remained bound to the cationic exchange column under these conditions. 36 mg of purified putative reductase protein (SEQ ID NO:2) were thus obtained.

Enzyme Characterization — Optimum temperature and pH were determined using standard (non-radioactive) assay conditions as described herein, except that the buffer concentration was changed to 19 mM and the protein following the Affi-Blue column chromatographic step was used (30 µl). Product formation analysis was investigated by reversed-phase HPLC analysis, with the following exceptions: for temperature optimum, incubations were performed at constant pH (7.0) with varying temperature (6 – 58 °C); for pH optimum, incubations were performed at constant temperature (30 °C) with varying pH (5.5 – 9.5). Initial velocity kinetics were analyzed by assaying the protein activity under standard (non-radioactive) conditions at pH 7.0, but with eleven different concentrations of lignan substrate (0.167 – 2.5 mM) and at 22 °C for 6 hours, while holding the NADPH concentration constant at 5 mM. To determine whether the 4*R*- or the 4*S*-hydride of the NADPH cofactor was utilized in the reduction catalyzed by the enzyme, radiochemical assays were performed with specifically labeled (*4R*-[^{3}H]NADPH or 4*S*-[^{3}H]NADPH) and analyzed for radiochemical incorporation into the 7-*O*-4′-(iso)dihydrodehydrodiconiferyl alcohol product

Radiochemical Assays for Dehydrodiconiferyl Alcohol Reductase Activity — Each 150 µl assay consisted of 19 mM MES:Bis-Tris Propane, pH 6.5, 20 µl of the protein solution at the corresponding stage of purity, 5 mM DTT, 2.5 mM (±)-dehydrodiconiferyl alcohols and 5 mM 4*R*-[^{3}H]-NADPH (14.2×10^3 kBq·mmol^{-1}). After 6 h incubation at 22 °C, the assay mixture was extracted with EtOAc (2 × 500 µl). The EtOAc soluble fraction was evaporated to dryness *in vacuo*, reconstituted in 100 µl of CH_3CN:3 % acetic acid (1:9) and subjected to reversed-phase HPLC as described herein, with both UV and radiochemical detection. One ml fractions were collected, with an aliquot (100 µl) of each removed for scintillation counting to determine the level of incorporation of ^{3}H into the assay products. Controls using denatured enzyme (5 min, 100 °C) or no (±)-dehydrodiconiferyl alcohol substrate were also performed.

Non-Radioactive Assays for Dehydrodiconiferyl Alcohol Reductase Activity — Each 150 µl assay consisted of 22 mM MES:Bis-Tris Propane, pH 6.5, 20

μl of the protein solution at the corresponding stage of purity, 5 mM DTT, 2 mM (±)-dehydrodiconiferyl alcohols and 4 mM NADPH. After 3 h incubation at 30 °C, the assay mixture was boiled for 3 min and centrifuged (17,000 ×g, 3 min), with an aliquot (125 μl) subjected to reversed-phase HPLC, after 16.6 μl of CH_3CN was added. Quantification of formation of 7-*O*-4′-dihydrodehydrodiconiferyl alcohol was measured using a standard curve, previously prepared. Controls using denatured enzyme (5 min, 100 °C) and no (±)-dehydrodiconiferyl alcohols as substrates were also performed.

HPLC Separation of Dehydrodiconiferyl Alcohol and its Reduced Products —Separation of dehydrodiconiferyl alcohol, dihydrodehydrodiconiferyl alcohol (7′,8′-allylic bond reduced), 7-*O*-4′-(iso)dihydrodehydrodiconiferyl alcohol and 7',8',7-O-4"-tetrahydrodehydrodiconiferyl alcohol was accomplished over a reversed-phase column (Symmetry Shield RP_8, 3.9 mm × 150 mm, Waters, Inc.) utilizing an acetonitrile: 3% acetic acid (in H_2O) solvent system as follows. The column was pre-equilibrated in CH_3CN:3 % acetic acid (A:B, 1:9). After injection of the sample to be analyzed, elution of the three compounds was achieved using the following elution profile: A:B (1:9) for 5 min, then a linear gradient over 30 min to A:B (25:75) and, finally, a linear gradient to 100% A over 25 min, with a flow rate of 1 $ml \cdot min^{-1}$.

*Assays for 7′,8′-*Dihydrodehydrodiconiferyl *Alcohol Reductase Activity* — Assays were performed, using both unlabeled and radiolabeled substrates as described above for (±)-dehydrodiconiferyl alcohol, but with 7′,8′-dihydrodehydrodiconiferyl alcohol added instead as substrate.

Since the reductase (SEQ ID NO:2) showed no ability to reduce pinoresinol or lariciresinol when expressed as a β-galactosidase fusion protein, native expression was attempted. The overexpression plasmid, pSBETa, contains the pET3a expression cassette, with two restriction sites available for sub-cloning (*Nde*I for native expression and *BamH*I for expression with a small fusion), as well as the *argU* gene (for production of the rare AGA-Arg $tRNA_{arg4}$) and kanamycin resistance (for high plasmid stability in liquid cultures). Since the putative reductase from *P. taeda* had no internal *Nde*I sites, sub-cloning into pSBETa was relatively straightforward. Two primers were used in PCR reactions to introduce *Nde*I sites at the start methionine (primer PT-ATG-NdeI) (SEQ ID NO:10) and in the 3′-end untranslated region (primer PT-Rev-NdeI) (SEQ ID NO:11) of the putative reductase (SEQ ID NO:2). The resulting PCR fragment was subcloned into a PCR cloning plasmid (pCRII, Invitrogen) and excised by cleavage with *Nde*I. The ~1 kb fragment containing the

putative reductase (SEQ ID NO:2) was gel purified and ligated into pSBETa, which had been previously digested with *Nde*I. The resulting pSBET construct containing the putative reductase (SEQ ID NO:2) was transformed into competent *E. coli* cells, purified and the expression region, containing the desired cDNA, was sequenced completely on both strands to verify that no mutations had been introduced during PCR.

Induction of the putative reductase (SEQ ID NO:2) in the pSBET over-expression system was accomplished by transforming B834(DE3) expression cells with the pSBET construct, and then growing four 1 liter cultures in LB broth supplemented with 50 $\mu g \cdot ml^{-1}$ kanamycin until they reached a density of $OD_{600} \cong 0.6$, at which time they were cooled to ~22 °C and inoculated with IPTG to a final concentration of 1 mM. After growth for 21 h at ~20-22°C, the cells were harvested by centrifugation.

Purification of the heterologously expressed, native enzyme (SEQ ID NO:2) to apparent homogeneity was accomplished in four chromatographic steps following cell lysis (as set forth herein) and cellular debris removal by centrifugation. The first step employed ammonium sulfate precipitation, with the desired protein pelleting in the 40-70 % saturation cut. This fraction of the protein (513 mg) was desalted into Buffer A and applied to an affinity (Affi Blue gel) column and eluted with a linear salt gradient. The resulting protein (118 mg) was again desalted into Buffer A and then subjected to consecutive anionic exchange (POROS 20 QE) and cationic exchange (POROS 20 SP) chromatography. Both were conducted under conditions (pH 6.8) where the desired protein did not bind, whereas the contaminating *E. coli* proteins did, and were thus removed. Additionally, after each step, the reductase (SEQ ID NO:2) was assayed for pinoresinol-lariciresinol reductase activity, but none was detected.

Moreover, even after the reductase (SEQ ID NO:2) was purified to near homogeneity, it was still unable to reduce either pinoresinol or lariciresinol. Since this putative reductase from *Pinus taeda* was closely aligned to the so-called isoflavone/pinoresinol-lariciresinol reductase "homologs", based on deduced amino acid sequence similarity, an alternative function was next considered. In this regard, since dehydrodiconiferyl alcohol contains a structure capable of forming a conjugated enone, and since *P. taeda* can contain significant levels of dehydrodiconiferyl alcohol, and metabolites thereof, we next evaluated the reductase (SEQ ID NO:2) for its ability to reduce dehydrodiconiferyl alcohol, either at the allylic position or at the

benzylic ether position, the latter being analogous to that catalyzed by pinoresinol-lariciresinol reductase (Gang, D.R., Fujita, M., Davin, L.B. & Lewis, N.G. *ACS Symp. Ser.* **697**:389-421 (1998).

Before such reactions could be analyzed, conditions for the HPLC separation of dehydrodiconiferyl alcohol, 7′,8′-dihydrodehydrodiconiferyl alcohol (allylic bond reduced), 7*O*4′-(iso)dihydrodehydrodiconiferyl alcohol (benzylic ether reduced), and 7′,8′,7*O*4′-tetrahydrodehydrodiconiferyl alcohol (both allylic bond and benzylic ether reduced) had to be developed. This was accomplished using a gradient acetonitrile:3% acetic acid (in H_2O) solvent system, as set forth *supra*. This solvent system gave near baseline separation of all four lignans, as shown in Figure 1A.

Assays for the putative reductase (SEQ ID NO:2) for dehydrodiconiferyl alcohol reductase activity (either allylic or benzylic ether) were performed as described herein. As shown in Figure 1B, the purified *P. taeda* reductase effects the reduction of the benzylic ether bond of dehydrodiconiferyl alcohol to convert it into 7-*O*-4′-(iso)dihydrodehydrodiconiferyl alcohol. The activity of the *P. taeda* putative reductase (SEQ ID NO:2) toward dehydrodiconiferyl alcohol at all steps of the purification scheme are listed in Table 1.

Table 1

Purification Scheme for the Benzylic Ether Reductase from *Pinus taeda* (SEQ ID NO:2)

Purification Step	Protein (mg)	Specific Activity ($nmol \cdot h^{-1} \cdot mg^{-1}$)
Lysate	1052	11
$(NH_4)_2SO_4$ precipitation	513	31
Affi-Blue	118	72
Anion Exchange (POROS-QE)	59	121
Cation Exchange (POROS-SP)	36	134

The authenticity of the substrate and the product were determined using LC-MS (see Figure 1D and 1E). Thus, the so-called isoflavone/pinoresinol-lariciresinol reductase "homologs" appear to be, in actuality, a phenylcoumaran benzylic ether reductase (PCBER) able to reduce dehydrodiconiferyl alcohol at the 7-O-4' position. As shown in Table 2, initial characterization of the enzyme (SEQ ID NO:2) revealed a pH optimum plateau from 6.5 to 7.0 and a temperature optimum at 49 °C. In addition, it is a Type A reductase, with transfer of the hydride of the

NADPH cofactor occurring almost exclusively from the 4*R*- position on the nicotinamide ring (Table 2).

Table 2

Properties of Benzylic Ether Reductase (SEQ ID NO:2) from *Pinus taeda*

	DDC	Substrate DDDC
K_M (mM)	0.61 ±0.03	1.95 ±0.1
V_{max} (nmol·h^{-1}·mg^{-1})	104.2 ±10.8	55.87 ±2.79
^{3}H-Isotope effect (V_{1H}/V_{3H})	14.4	4.0
Hydride abstraction from NADPH	4*pro-R* (>99%)	–
pH optimum	6.4-7.0	–
Temp. optimum (°C)	49	–

When 7′,8′-dihydrodehydrodiconiferyl alcohol was tested as the substrate, reduction of the benzylic ether was also accomplished, as evidenced by the formation of tetrahydrodehydrodiconiferyl alcohol (see Figure 1C and 1F).

Initial velocity studies were carried out in order to determine whether the *P. taeda* benzylic ether reductase (SEQ ID NO:2) preferred one of these substrates over the other. These results are listed in Table 2. As can be seen, the specific activity (V_{max}) is 2-fold higher for dehydrodiconiferyl alcohol as substrate than for 7′,8′-dihydrodehydrodiconiferyl alcohol. In addition, the K_M for dehydrodiconiferyl alcohol reduction is 3 times lower than the K_M for 7′,8′-dihydrodehydrodiconiferyl alcohol reduction. Interestingly, the kinetic isotope effect (V_{1H}/V_{3H}) observed (which indicates that the hydride transfer step of the reaction contributes to the rate limiting step[s] of the reaction) is significantly higher for dehydrodiconiferyl alcohol reduction. When dehydrodiconiferyl alcohol is the substrate, the hydride transfer contributes significantly (as indicated by a high kinetic isotope effect, V_{1H}/V_{3H}=14.4) to the rate limiting step of the enzyme catalyzed reaction. When 7′,8′-dihydrodehydrodiconiferyl alcohol is given as subtrate, however, the kinetic isotope effect is reduced to V_{1H}/V_{3H}=4.0, indicating that the hydride transfer is no longer contributing significantly to the rate limiting step of the reaction. This is also supported by the significant increase in the K_M value for 7′,8′-dihydrodehydrodiconiferyl alcohol reduction, indicating that binding of the substrate is much less efficient for this lignan than for dehydrodiconiferyl alcohol.

These results indicate that the *P. taeda* benzylic ether reductase (SEQ ID NO:2) has a significantly higher affinity for dehydrodiconiferyl alcohol as substrate.

EXAMPLE 4

Cloning of Dehydrodiconiferyl Alcohol Reductase cDNAs (SEQ ID NO:3) and (SEQ ID NO:5) from *Cryptomeria japonica*

Cryptomeria japonica cDNA Library Synthesis — Total RNA (200 μg/g fresh weight) was obtained using the method of Dong and Dunsten (Dong, J.Z. & Dunstan, D.I. *Plant Cell Reports* **15** (1996)) from leaves of greenhouse-grown *C. japonica* trees. A *C. japonica* cDNA library was constructed using 5 μg of purified poly(A)$^+$ mRNA (Oligotex-dT™ Suspension, Qiagen) with the ZAP-cDNA® synthesis kit, the Uni ZAP™ XR vector, and the Gigapack® II Gold packaging extract (Stratagene), with a titer of 2.2 X 10^6 pfu for the primary library. The amplified library (8.3 X 10^9 pfu /ml; 175 ml total) was used for screening (Dinkova-Kostova, A.T., *et al. J. Biol. Chem.* **271**, 29473-29482 (1996)).

DNA Probe Synthesis — The pSBETa plasmid containing the *P. taeda* dehydrodiconiferyl alcohol reductase cDNA (5 ng) (SEQ ID NO:1) was used as the template in five 100 μl PCR reactions (10 mM Tris-HCl [pH 9.0], 50 mM KCl, 0.1% Triton X-100, 2.5 mM $MgCl_2$, 0.2 mM each dNTP and 2.5 units *Taq* DNA polymerase) with primers Cj-PCBERNT (100 pmol) (SEQ ID NO:12) and Cj-PCBERCT (100 pmol) (SEQ ID NO:13). The PCR amplifications were carried out in a thermocycler as follows: 35 cycles of 1 min at 94°C, 2 min at 50°C and 3 min at 72°C; with 5 min at 72°C and an indefinite hold at 4°C after the final cycle. Single primer, template-only and primer-only reactions were performed as controls.

The 5 reactions were concentrated (Microcon 30, Amicon Inc.) and washed with TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA; 2 x 200 μl), with the PCR product subsequently recovered in TE buffer (2 x 50 μl). The PCR product (~980 bp band) was resolved in a preparative 1.0% agarose gel and purified from the agarose using the GENECLEAN II® kit (BIO 101 Inc.). The gel-purified PCR product (40 ng) was used with Pharmacia's ^{T7}QuickPrime® kit and [α-^{32}P]dCTP, according to kit instructions, to produce a radiolabeled probe (in 0.1 ml), which was purified over BioSpin 6 columns (Bio-Rad) and added to carrier DNA (0.5 mg/ml sheared salmon sperm DNA [Sigma], 0.9 ml).

Library Screening - 300,000 pfu of *C. japonica* amplified cDNA library were plated for primary screening, according to Stratagene's instructions.

Plaques were blotted onto Magna Nylon membrane circles (Micron Separations Inc.), which were then allowed to air dry. The membranes were placed between two layers of Whatman® 3MM Chr paper. cDNA Library phage DNA was fixed to the membranes and denatured in one step by autoclaving for 2 min at 100 °C with fast exhaust. The membranes were washed for 30 min at 37°C in 6X standard saline citrate (SSC) and 0.1% SDS and prehybridized for 5 h with gentle shaking at 49°C in preheated 6X SSC, 0.5% SDS and 5X Denhardt's reagent (hybridization solution, 300 ml) in a crystallization dish (190 x 75 mm). The [^{32}P]radiolabeled probe was denatured (boiling, 10 min), quickly cooled (ice, 15 min) and added to a preheated fresh hybridization solution (60 ml, 49°C) in a crystallization dish (150 x 75 mm). The prehybridized membranes were next added to this dish, which was then covered with plastic wrap. Hybridization was performed for 18 h at 49°C with gentle shaking. The membranes were washed in 4X SSC and 0.5% SDS for 5 min at room temperature, transferred to 2X SSC and 0.5% SDS (at room temperature) and incubated at 49°C for 20 min with gentle shaking, wrapped with plastic wrap to prevent drying and finally exposed to Kodak X-OMAT AR film for 24 h at -80°C with intensifying screens. Twenty positive plaques were purified through two more rounds of screening with hybridization conditions as above, and two were found to encode the expected enzyme.

In vivo Excision and Sequencing of pCj-PCBER1 and pCj-PCBER2 Phagemids - Purified cDNA clones were rescued from the phage following Stratagene's *in vivo* excision protocol. Both strands of the 2 different cDNAs (pCj-PCBER1 (SEQ ID NO:3) and pCj-PCBER2 (SEQ ID NO:5)) were completely sequenced using overlapping sequencing primers.

EXAMPLE 5

Properties of Presently Preferred Dehydrodiconiferyl Alcohol Benzylic Ether Reductase Proteins of the Present Invention

Presently preferred dehydrodiconiferyl alcohol benzylic ether reductase proteins of the present invention are NADPH dependent reductases having a molecular weight (as determined by SDS PAGE) of from about 33 kDa to about 34 kDa, a pH optimum in the range of from about 6.0 to about 7.0, and a pI value (computer generated) of from about 6.0 to about 7.0. Additionally, presently most preferred dehydrodiconiferyl alcohol benzylic ether reductase proteins of the present invention have a Vmax of 104.2 ± 10.8 nmol/h/mg, and a Km of 0.61 ± 0.03 mM.

Presently preferred dehydrodiconiferyl alcohol benzylic ether reductase proteins of the invention (and nucleic acids that encode them) are from gymnosperm or angiosperm plant species.

While the preferred embodiment of the invention has been illustrated and described, it will be appreciated that various changes can be made therein without departing from the spirit and scope of the invention.

The embodiments of the invention in which an exclusive property or privilege is claimed are defined as follows:

1. An isolated nucleic acid molecule encoding a dehydrodiconiferyl alcohol benzylic ether reductase protein.

2. An isolated nucleic acid molecule of Claim 1 encoding a gymnosperm dehydrodiconiferyl alcohol benzylic ether reductase protein.

3. An isolated nucleic acid molecule of Claim 1 encoding a dehydrodiconiferyl alcohol benzylic ether reductase protein from the genus *Pinus*.

4. An isolated nucleic acid molecule of Claim 3 encoding a dehydrodiconiferyl alcohol benzylic ether reductase protein from *Pinus taeda*.

5. An isolated nucleic acid molecule of Claim 4 comprising the nucleotide sequence of SEQ ID NO:1.

6. An isolated nucleic acid molecule of Claim 1 encoding a protein consisting of the amino acid sequence of SEQ ID NO:2.

7. An isolated nucleic acid molecule of Claim 1 encoding a dehydrodiconiferyl alcohol benzylic ether reductase protein from the genus *Cryptomeria*.

8. An isolated nucleic acid molecule of Claim 7 encoding a dehydrodiconiferyl alcohol benzylic ether reductase protein from *Cryptomeria japonica*.

9. An isolated nucleic acid molecule of Claim 8 comprising the nucleotide sequence of SEQ ID NO:3.

10. An isolated nucleic acid molecule of Claim 8 comprising the nucleotide sequence of SEQ ID NO:5.

11. An isolated nucleic acid molecule of Claim 1 encoding a protein consisting of the amino acid sequence of SEQ ID NO:4.

12. An isolated nucleic acid molecule of Claim 1 encoding a protein consisting of the amino acid sequence of SEQ ID NO:6.

13. An isolated dehydrodiconiferyl alcohol benzylic ether reductase protein.

14. An isolated angiosperm dehydrodiconiferyl alcohol benzylic ether reductase protein of Claim 13.

15. An isolated gymnosperm dehydrodiconiferyl alcohol benzylic ether reductase protein of Claim 13.

16. An isolated *Pinus* dehydrodiconiferyl alcohol benzylic ether reductase protein of Claim 13.

17. An isolated *Pinus taeda* dehydrodiconiferyl alcohol benzylic ether reductase protein of Claim 13.

18. An isolated dehydrodiconiferyl alcohol benzylic ether reductase protein of Claim 13, said protein comprising the amino acid sequence of SEQ ID NO:2.

19. An isolated *Cryptomeria* dehydrodiconiferyl alcohol benzylic ether reductase protein of Claim 13.

20. An isolated *Cryptomeria japonica* dehydrodiconiferyl alcohol benzylic ether reductase protein of Claim 13.

21. An isolated dehydrodiconiferyl alcohol benzylic ether reductase protein of Claim 13, said protein comprising the amino acid sequence of SEQ ID NO:4.

22. An isolated dehydrodiconiferyl alcohol benzylic ether reductase protein of Claim 13, said protein comprising the amino acid sequence of SEQ ID NO:6.

23. A replicable expression vector comprising a nucleotide sequence encoding a dehydrodiconiferyl alcohol benzylic ether reductase protein.

24. A replicable expression vector of Claim 23 comprising a nucleotide sequence encoding a gymnosperm dehydrodiconiferyl alcohol benzylic ether reductase protein.

25. A replicable expression vector of Claim 23 comprising a nucleotide sequence encoding an angiosperm dehydrodiconiferyl alcohol benzylic ether reductase protein.

26. A replicable expression vector of Claim 23 comprising a nucleotide sequence encoding a *Pinus* dehydrodiconiferyl alcohol benzylic ether reductase protein.

27. A replicable expression vector of Claim 23 comprising a nucleotide sequence encoding a *Cryptomeria* dehydrodiconiferyl alcohol benzylic ether reductase protein.

28. A method of enhancing the expression of dehydrodiconiferyl alcohol benzylic ether reductase protein in a suitable host cell comprising introducing into the host cell an expression vector that comprises a nucleotide sequence encoding a dehydrodiconiferyl alcohol benzylic ether reductase protein.

29. The method of Claim 28 wherein the host cell is a plant cell.

30. A method of modifying the expression of dehydrodiconiferyl alcohol benzylic ether reductase protein in a suitable host cell comprising introducing into the host cell vector comprising a nucleotide sequence that expresses an RNA that is capable of hybridizing to the cDNA molecule set forth in SEQ ID NO:1.

31. The method of Claim 30 wherein the host cell is a plant cell.

32. A method of modifying the expression of dehydrodiconiferyl alcohol benzylic ether reductase protein in a suitable host cell comprising introducing into the host cell vector comprising a nucleotide sequence that expresses an RNA that is capable of hybridizing to the cDNA molecule set forth in SEQ ID NO:3.

33. The method of Claim 32 wherein the host cell is a plant cell.

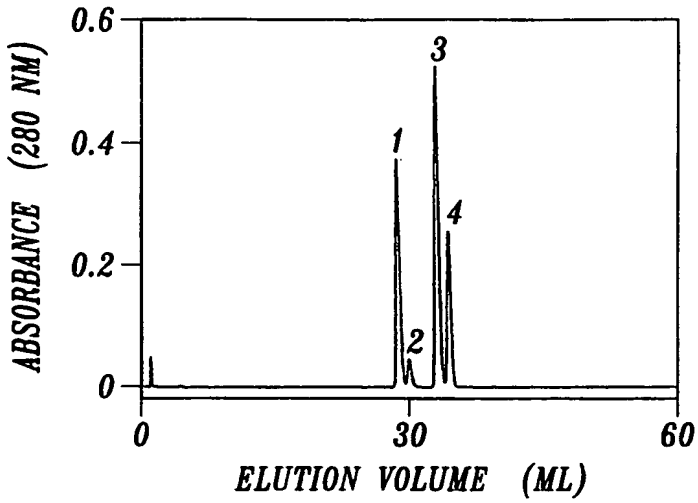

Fig. 1A

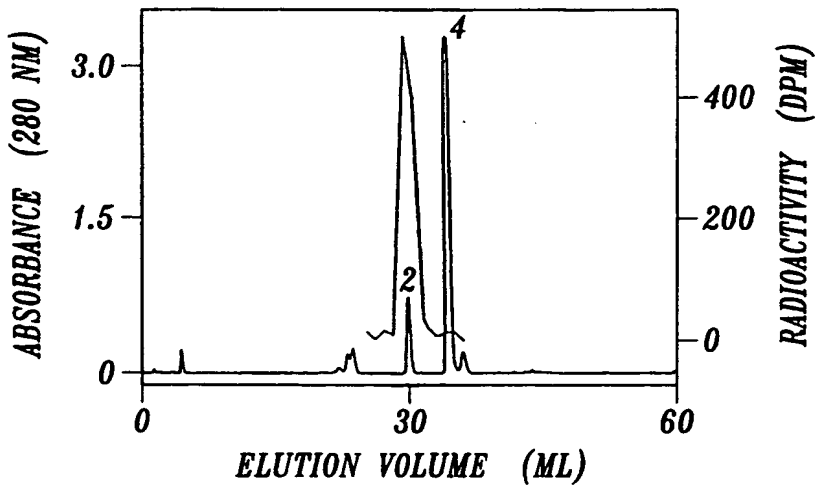

Fig. 1B

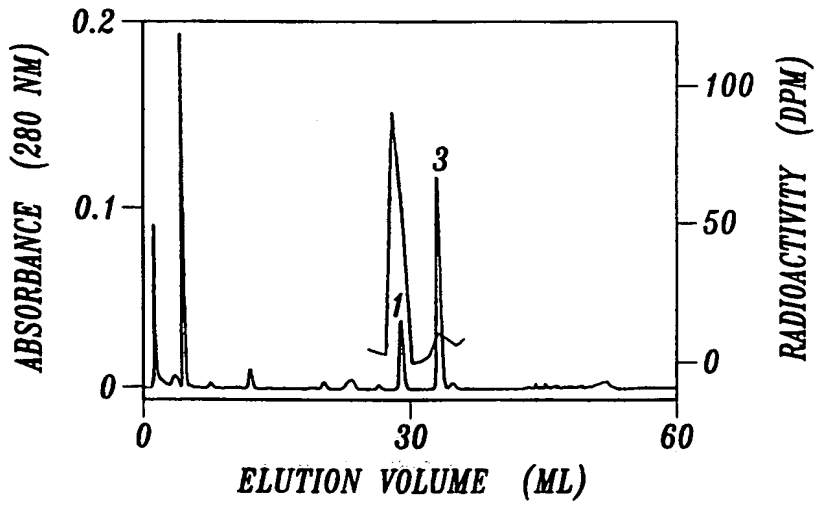

Fig. 1C

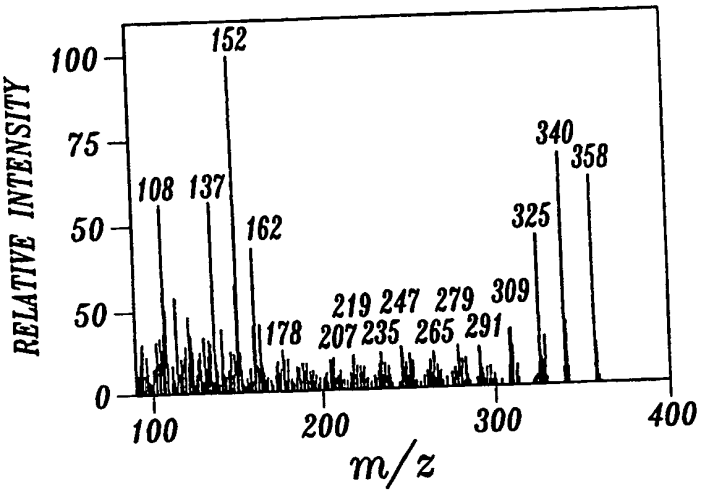

Fig. 1D

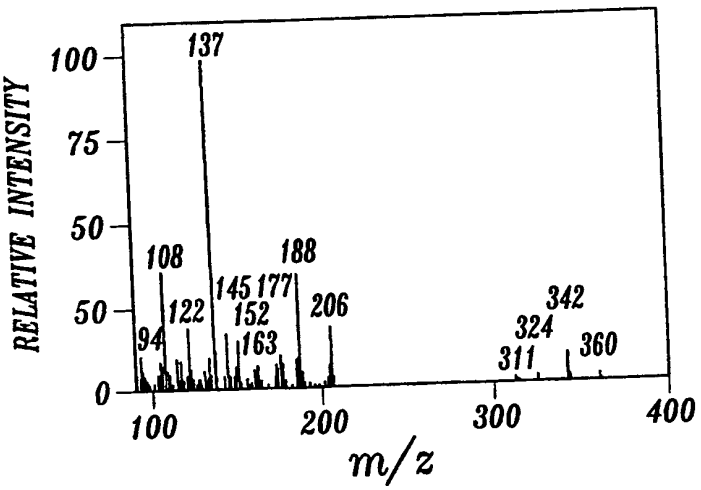

Fig. 1E

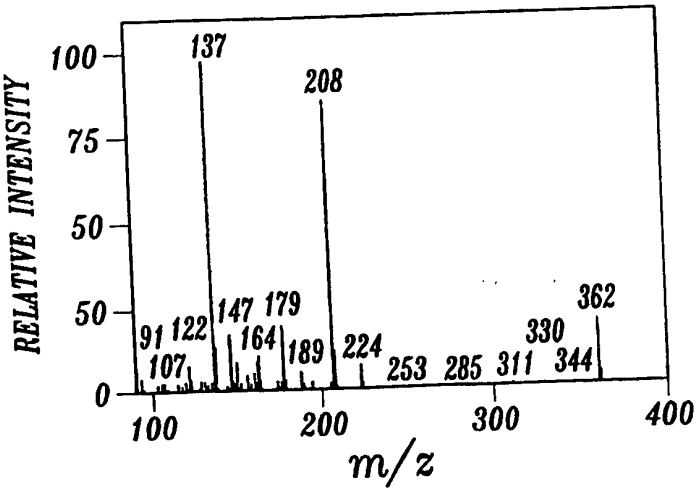

Fig. 1F

SEQUENCE LISTING

<110> Lewis, Norman G
Kasahara, Hiroyuki
Gang, David R
Davin, Laurence B

<120> Recombinant Dehydrodiconiferyl Alcohol Benzylic Ether Reductase and Methods of Use

<130> wsur114203

<140>
<141>

<150> 60/094,012
<151> 1998-07-24

<160> 13

<170> PatentIn Ver. 2.0

<210> 1
<211> 926
<212> DNA
<213> Pinus taeda

<220>
<221> CDS
<222> (1)..(924)

<400> 1

```
atg gga agc agg agc agg ata ctc cta att ggc gca aca gga tac att   48
Met Gly Ser Arg Ser Arg Ile Leu Leu Ile Gly Ala Thr Gly Tyr Ile
  1               5                  10                  15

ggt cgc cat gtt gcc aag gct agc ctt gat ctc ggc cat ccc acc ttc   96
Gly Arg His Val Ala Lys Ala Ser Leu Asp Leu Gly His Pro Thr Phe
             20                  25                  30

ctt ctg gtt aga gag tcc act gct tct tct aat tct gag aaa gcc cag   144
Leu Leu Val Arg Glu Ser Thr Ala Ser Ser Asn Ser Glu Lys Ala Gln
         35                  40                  45

ctc ctg gaa tcc ttc aag gcc tct ggt gct aat ata gtc cat gga tcc   192
Leu Leu Glu Ser Phe Lys Ala Ser Gly Ala Asn Ile Val His Gly Ser
     50                  55                  60

ata gat gat cat gca agc ctt gtg gag gca gtg aag aat gtg gat gta   240
Ile Asp Asp His Ala Ser Leu Val Glu Ala Val Lys Asn Val Asp Val
 65                  70                  75                  80

gta atc tcc aca gtt gga tca cta cag ata gag agc cag gtc aat att   288
Val Ile Ser Thr Val Gly Ser Leu Gln Ile Glu Ser Gln Val Asn Ile
                 85                  90                  95

atc aag gct att aaa gaa att gga acc gtc aag agg ttt ttt cca tct   336
Ile Lys Ala Ile Lys Glu Ile Gly Thr Val Lys Arg Phe Phe Pro Ser
```

```
                100                 105                 110
gag ttc ggg aat gat gtt gat aac gtc cat gca gtg gag cct gca aag   384
Glu Phe Gly Asn Asp Val Asp Asn Val His Ala Val Glu Pro Ala Lys
        115                 120                 125

aat gtg ttt gag gtg aaa gcc aaa gtc cgt agg gca atc gaa gca gag   432
Asn Val Phe Glu Val Lys Ala Lys Val Arg Arg Ala Ile Glu Ala Glu
    130                 135                 140

ggt att cct tat aca tac gtc tct agc aac tgt ttt gca ggg tat ttc   480
Gly Ile Pro Tyr Thr Tyr Val Ser Ser Asn Cys Phe Ala Gly Tyr Phe
145                 150                 155                 160

ctg cga agc ctc gca cag gct ggc cta aca gct cct cca aga gat aaa   528
Leu Arg Ser Leu Ala Gln Ala Gly Leu Thr Ala Pro Pro Arg Asp Lys
                165                 170                 175

gtt gtc att ctt gga gat gga aat gcc aga gtt gtc ttt gta aaa gag   576
Val Val Ile Leu Gly Asp Gly Asn Ala Arg Val Val Phe Val Lys Glu
            180                 185                 190

gaa gac att gga aca ttt aca atc aag gca gtg gac gac ccc aga acg   624
Glu Asp Ile Gly Thr Phe Thr Ile Lys Ala Val Asp Asp Pro Arg Thr
        195                 200                 205

ttg aac aag acc tta tac ttg agg ctt cct gcc aat act ctg tct tta   672
Leu Asn Lys Thr Leu Tyr Leu Arg Leu Pro Ala Asn Thr Leu Ser Leu
    210                 215                 220

aat gag ctt gtg gct ctc tgg gag aag aag att gat aag act ctg gag   720
Asn Glu Leu Val Ala Leu Trp Glu Lys Lys Ile Asp Lys Thr Leu Glu
225                 230                 235                 240

aag gcc tac gtg ccc gag gag gag gtt ctt aaa tta atc gca gat aca   768
Lys Ala Tyr Val Pro Glu Glu Glu Val Leu Lys Leu Ile Ala Asp Thr
                245                 250                 255

cca ttc cca gct aat att agc ata gca att agt cat tct atc ttc gtg   816
Pro Phe Pro Ala Asn Ile Ser Ile Ala Ile Ser His Ser Ile Phe Val
            260                 265                 270

aaa gga gat caa aca aat ttt gaa att gga cct gct ggt gta gag gct   864
Lys Gly Asp Gln Thr Asn Phe Glu Ile Gly Pro Ala Gly Val Glu Ala
        275                 280                 285

agt cag ctg tac cct gat gtg aaa tac acc acc gtc gac gag tac ctg   912
Ser Gln Leu Tyr Pro Asp Val Lys Tyr Thr Thr Val Asp Glu Tyr Leu
    290                 295                 300

agc aat ttt gtg tg                                               926
Ser Asn Phe Val
305
```

<210> 2
<211> 308
<212> PRT

<213> Pinus taeda

<400> 2

```
Met Gly Ser Arg Ser Arg Ile Leu Leu Ile Gly Ala Thr Gly Tyr Ile
  1               5                  10                  15

Gly Arg His Val Ala Lys Ala Ser Leu Asp Leu Gly His Pro Thr Phe
             20                  25                  30

Leu Leu Val Arg Glu Ser Thr Ala Ser Ser Asn Ser Glu Lys Ala Gln
         35                  40                  45

Leu Leu Glu Ser Phe Lys Ala Ser Gly Ala Asn Ile Val His Gly Ser
     50                  55                  60

Ile Asp Asp His Ala Ser Leu Val Glu Ala Val Lys Asn Val Asp Val
 65                  70                  75                  80

Val Ile Ser Thr Val Gly Ser Leu Gln Ile Glu Ser Gln Val Asn Ile
                 85                  90                  95

Ile Lys Ala Ile Lys Glu Ile Gly Thr Val Lys Arg Phe Phe Pro Ser
            100                 105                 110

Glu Phe Gly Asn Asp Val Asp Asn Val His Ala Val Glu Pro Ala Lys
        115                 120                 125

Asn Val Phe Glu Val Lys Ala Lys Val Arg Arg Ala Ile Glu Ala Glu
    130                 135                 140

Gly Ile Pro Tyr Thr Tyr Val Ser Ser Asn Cys Phe Ala Gly Tyr Phe
145                 150                 155                 160

Leu Arg Ser Leu Ala Gln Ala Gly Leu Thr Ala Pro Pro Arg Asp Lys
                165                 170                 175

Val Val Ile Leu Gly Asp Gly Asn Ala Arg Val Val Phe Val Lys Glu
            180                 185                 190

Glu Asp Ile Gly Thr Phe Thr Ile Lys Ala Val Asp Asp Pro Arg Thr
        195                 200                 205

Leu Asn Lys Thr Leu Tyr Leu Arg Leu Pro Ala Asn Thr Leu Ser Leu
    210                 215                 220

Asn Glu Leu Val Ala Leu Trp Glu Lys Lys Ile Asp Lys Thr Leu Glu
225                 230                 235                 240

Lys Ala Tyr Val Pro Glu Glu Glu Val Leu Lys Leu Ile Ala Asp Thr
                245                 250                 255

Pro Phe Pro Ala Asn Ile Ser Ile Ala Ile Ser His Ser Ile Phe Val
            260                 265                 270

Lys Gly Asp Gln Thr Asn Phe Glu Ile Gly Pro Ala Gly Val Glu Ala
        275                 280                 285

Ser Gln Leu Tyr Pro Asp Val Lys Tyr Thr Thr Val Asp Glu Tyr Leu
```

290 295 300

Ser Asn Phe Val
305

<210> 3
<211> 918
<212> DNA
<213> Cryptomeria japonica

<220>
<221> CDS
<222> (1)..(918)

<400> 3

```
atg gga gga agt agg gtt ttg ata ata gga gga act ggt tac att ggt   48
Met Gly Gly Ser Arg Val Leu Ile Ile Gly Gly Thr Gly Tyr Ile Gly
  1               5                  10                  15

aga cat gtg acc aat gca agc ctt gcc caa ggt cac ccc act ttc ctc   96
Arg His Val Thr Asn Ala Ser Leu Ala Gln Gly His Pro Thr Phe Leu
            20                  25                  30

ctt gtt aga gaa atc aca cct tcc aac cct gaa aag gcc cag ctg cta   144
Leu Val Arg Glu Ile Thr Pro Ser Asn Pro Glu Lys Ala Gln Leu Leu
        35                  40                  45

gag tcc ttc aca tcc aaa gga gca act ctg gtt caa ggg tcc ata gat   192
Glu Ser Phe Thr Ser Lys Gly Ala Thr Leu Val Gln Gly Ser Ile Asp
    50                  55                  60

gat cat gca agc ctt gta gca gca ctg aag aag gtg gat gta gtc att   240
Asp His Ala Ser Leu Val Ala Ala Leu Lys Lys Val Asp Val Val Ile
65                  70                  75                  80

tca aca ttg ggt gca cca cag ata gca gat cag ttc aat cta atc aag   288
Ser Thr Leu Gly Ala Pro Gln Ile Ala Asp Gln Phe Asn Leu Ile Lys
                85                  90                  95

gct atc aaa gaa gtt gga act att aag agg ttt ttc cca tca gag ttt   336
Ala Ile Lys Glu Val Gly Thr Ile Lys Arg Phe Phe Pro Ser Glu Phe
            100                 105                 110

ggg aat gat gtt gat aag cac cat gca gtg gag cct atg aag agt atg   384
Gly Asn Asp Val Asp Lys His His Ala Val Glu Pro Met Lys Ser Met
        115                 120                 125

ttt gac ttg aaa atc aaa ctt cgc cgg act att gag gca gaa ggc att   432
Phe Asp Leu Lys Ile Lys Leu Arg Arg Thr Ile Glu Ala Glu Gly Ile
    130                 135                 140

cca cat aca tat gtt gtt cca cat tgt ttt gca ggc tac ttt ctt act   480
Pro His Thr Tyr Val Val Pro His Cys Phe Ala Gly Tyr Phe Leu Thr
145                 150                 155                 160

aat ctt gca cag ctt ggg cta gct gcc ccg cca agg gat aaa att gtt   528
Asn Leu Ala Gln Leu Gly Leu Ala Ala Pro Pro Arg Asp Lys Ile Val
```

```
                165                 170                 175

att tat gga gat gga act acc aaa gct gta tat atg aag gaa gaa gat   576
Ile Tyr Gly Asp Gly Thr Thr Lys Ala Val Tyr Met Lys Glu Glu Asp
            180                 185                 190

att ggg aca ttc aca atc aag gca gtg gat gat cca agg acc ttg aac   624
Ile Gly Thr Phe Thr Ile Lys Ala Val Asp Asp Pro Arg Thr Leu Asn
        195                 200                 205

aag acc ctc tac ttg aag cct cct gcc aat act atc tcc acc aat gat   672
Lys Thr Leu Tyr Leu Lys Pro Pro Ala Asn Thr Ile Ser Thr Asn Asp
    210                 215                 220

ctc gta gcc ctt tgg gag gca aag ata ggg aag act ttg gag aag gtc   720
Leu Val Ala Leu Trp Glu Ala Lys Ile Gly Lys Thr Leu Glu Lys Val
225                 230                 235                 240

tat cta tct gag gag cag gtt ctt aaa cta ctc caa gat aca cca ttt   768
Tyr Leu Ser Glu Glu Gln Val Leu Lys Leu Leu Gln Asp Thr Pro Phe
                245                 250                 255

cca ggg act ttt atg gta tcc att ttt cac acc att tat gtg aaa gga   816
Pro Gly Thr Phe Met Val Ser Ile Phe His Thr Ile Tyr Val Lys Gly
            260                 265                 270

gat caa acc aat ttt cag att gga cct gat ggg gtt gag gcc agt gcg   864
Asp Gln Thr Asn Phe Gln Ile Gly Pro Asp Gly Val Glu Ala Ser Ala
        275                 280                 285

ctg tac cca gat gtg aaa tat aca act gtt gaa gag tac atc agt gca   912
Leu Tyr Pro Asp Val Lys Tyr Thr Thr Val Glu Glu Tyr Ile Ser Ala
    290                 295                 300

ttt gta                                                           918
Phe Val
305
```

<210> 4
<211> 306
<212> PRT
<213> Cryptomeria japonica

<400> 4

```
Met Gly Gly Ser Arg Val Leu Ile Ile Gly Gly Thr Gly Tyr Ile Gly
  1               5                  10                  15

Arg His Val Thr Asn Ala Ser Leu Ala Gln Gly His Pro Thr Phe Leu
             20                  25                  30

Leu Val Arg Glu Ile Thr Pro Ser Asn Pro Glu Lys Ala Gln Leu Leu
         35                  40                  45

Glu Ser Phe Thr Ser Lys Gly Ala Thr Leu Val Gln Gly Ser Ile Asp
     50                  55                  60

Asp His Ala Ser Leu Val Ala Ala Leu Lys Lys Val Asp Val Val Ile
```

65 70 75 80

Ser Thr Leu Gly Ala Pro Gln Ile Ala Asp Gln Phe Asn Leu Ile Lys
85 90 95

Ala Ile Lys Glu Val Gly Thr Ile Lys Arg Phe Phe Pro Ser Glu Phe
100 105 110

Gly Asn Asp Val Asp Lys His His Ala Val Glu Pro Met Lys Ser Met
115 120 125

Phe Asp Leu Lys Ile Lys Leu Arg Arg Thr Ile Glu Ala Glu Gly Ile
130 135 140

Pro His Thr Tyr Val Val Pro His Cys Phe Ala Gly Tyr Phe Leu Thr
145 150 155 160

Asn Leu Ala Gln Leu Gly Leu Ala Ala Pro Pro Arg Asp Lys Ile Val
165 170 175

Ile Tyr Gly Asp Gly Thr Thr Lys Ala Val Tyr Met Lys Glu Glu Asp
180 185 190

Ile Gly Thr Phe Thr Ile Lys Ala Val Asp Asp Pro Arg Thr Leu Asn
195 200 205

Lys Thr Leu Tyr Leu Lys Pro Pro Ala Asn Thr Ile Ser Thr Asn Asp
210 215 220

Leu Val Ala Leu Trp Glu Ala Lys Ile Gly Lys Thr Leu Glu Lys Val
225 230 235 240

Tyr Leu Ser Glu Glu Gln Val Leu Lys Leu Leu Gln Asp Thr Pro Phe
245 250 255

Pro Gly Thr Phe Met Val Ser Ile Phe His Thr Ile Tyr Val Lys Gly
260 265 270

Asp Gln Thr Asn Phe Gln Ile Gly Pro Asp Gly Val Glu Ala Ser Ala
275 280 285

Leu Tyr Pro Asp Val Lys Tyr Thr Thr Val Glu Glu Tyr Ile Ser Ala
290 295 300

Phe Val
305

<210> 5
<211> 963
<212> DNA
<213> Cryptomeria japonica

<220>
<221> CDS
<222> (1)..(918)

<400> 5

```
atg gga ggc agc aga att ctg ttg att gga gga act gga tac att ggt   48
Met Gly Gly Ser Arg Ile Leu Leu Ile Gly Gly Thr Gly Tyr Ile Gly
  1               5                  10                  15

cga cat gtg gca aaa gcc agc att gct ctt ggc cat ccc act tac ctt   96
Arg His Val Ala Lys Ala Ser Ile Ala Leu Gly His Pro Thr Tyr Leu
             20                  25                  30

ctg gtc aga gag tcc acg gct tca aat cct gtt aaa gct gag ctt ctg  144
Leu Val Arg Glu Ser Thr Ala Ser Asn Pro Val Lys Ala Glu Leu Leu
         35                  40                  45

gag tcc ttc aag gca tca ggt ggt caa ata gtg cat gga tct ctg ggt  192
Glu Ser Phe Lys Ala Ser Gly Gly Gln Ile Val His Gly Ser Leu Gly
     50                  55                  60

gat cat aca agt ctt gtg gag gca atc aag aag gtt gat gtg gtt ata  240
Asp His Thr Ser Leu Val Glu Ala Ile Lys Lys Val Asp Val Val Ile
 65                  70                  75                  80

tcc aca gtg gga ggg gcg cag ata atg gat cag tta aac att atc aag  288
Ser Thr Val Gly Gly Ala Gln Ile Met Asp Gln Leu Asn Ile Ile Lys
                 85                  90                  95

gca atc aaa gag gtt gga acc att aag aga ttt gtt cca tct gag ttc  336
Ala Ile Lys Glu Val Gly Thr Ile Lys Arg Phe Val Pro Ser Glu Phe
            100                 105                 110

gga aat gat gtt gac aaa gtc cat gca gtg gag cct gct aag acc atg  384
Gly Asn Asp Val Asp Lys Val His Ala Val Glu Pro Ala Lys Thr Met
        115                 120                 125

ttt ggt aat aaa gcc aaa atc cgc agg gcc att gag gca gaa ggc att  432
Phe Gly Asn Lys Ala Lys Ile Arg Arg Ala Ile Glu Ala Glu Gly Ile
    130                 135                 140

ccc tac acg tat gtt tcc agc aat tta ttt gca ggt tat ttt ttg cca  480
Pro Tyr Thr Tyr Val Ser Ser Asn Leu Phe Ala Gly Tyr Phe Leu Pro
145                 150                 155                 160

aat ttg ggg caa cca ggc ctc tct gcc ccg cca agg gat aaa ctt gtt  528
Asn Leu Gly Gln Pro Gly Leu Ser Ala Pro Pro Arg Asp Lys Leu Val
                165                 170                 175

att tta ggg gat ggc aat gcc aaa gtt gtg tat gta aaa gag gag gat  576
Ile Leu Gly Asp Gly Asn Ala Lys Val Val Tyr Val Lys Glu Glu Asp
            180                 185                 190

att ggc acc ttc acc atc aag gca gtg gat gac cca aga act ctg aac  624
Ile Gly Thr Phe Thr Ile Lys Ala Val Asp Asp Pro Arg Thr Leu Asn
        195                 200                 205

aag acc ctg tat ttg agg ctt cct tcc aat acc tat tct ttt aat gag  672
Lys Thr Leu Tyr Leu Arg Leu Pro Ser Asn Thr Tyr Ser Phe Asn Glu
    210                 215                 220

ctg gca gcc ttg tgg gag acg aag ata gga aag act ctg gag aag gtc  720
Leu Ala Ala Leu Trp Glu Thr Lys Ile Gly Lys Thr Leu Glu Lys Val
```

225 230 235 240

tat gtc cct gag gaa gct gtc ctc aaa aaa att gaa gag aca cca ttc 768
Tyr Val Pro Glu Glu Ala Val Leu Lys Lys Ile Glu Glu Thr Pro Phe
245 250 255

cca gcc agt att ggt att tca att ggt tac tct act ttt gtg aaa ggg 816
Pro Ala Ser Ile Gly Ile Ser Ile Gly Tyr Ser Thr Phe Val Lys Gly
260 265 270

gac caa acc aat ttt gaa att gga cct gat ggt gtg gag ggc agc caa 864
Asp Gln Thr Asn Phe Glu Ile Gly Pro Asp Gly Val Glu Gly Ser Gln
275 280 285

ctc tac cca gat gtg aag tac acc act gtt gaa gag tat ctt gac cag 912
Leu Tyr Pro Asp Val Lys Tyr Thr Thr Val Glu Glu Tyr Leu Asp Gln
290 295 300

ttt gtt tgactaagaa ttgaaggtga caatatacta atatagtggg atgta 963
Phe Val
305

<210> 6
<211> 306
<212> PRT
<213> Cryptomeria japonica

<400> 6
Met Gly Gly Ser Arg Ile Leu Leu Ile Gly Gly Thr Gly Tyr Ile Gly
1 5 10 15

Arg His Val Ala Lys Ala Ser Ile Ala Leu Gly His Pro Thr Tyr Leu
20 25 30

Leu Val Arg Glu Ser Thr Ala Ser Asn Pro Val Lys Ala Glu Leu Leu
35 40 45

Glu Ser Phe Lys Ala Ser Gly Gly Gln Ile Val His Gly Ser Leu Gly
50 55 60

Asp His Thr Ser Leu Val Glu Ala Ile Lys Lys Val Asp Val Val Ile
65 70 75 80

Ser Thr Val Gly Gly Ala Gln Ile Met Asp Gln Leu Asn Ile Ile Lys
85 90 95

Ala Ile Lys Glu Val Gly Thr Ile Lys Arg Phe Val Pro Ser Glu Phe
100 105 110

Gly Asn Asp Val Asp Lys Val His Ala Val Glu Pro Ala Lys Thr Met
115 120 125

Phe Gly Asn Lys Ala Lys Ile Arg Arg Ala Ile Glu Ala Glu Gly Ile
130 135 140

Pro Tyr Thr Tyr Val Ser Ser Asn Leu Phe Ala Gly Tyr Phe Leu Pro
145 150 155 160

Asn Leu Gly Gln Pro Gly Leu Ser Ala Pro Pro Arg Asp Lys Leu Val
165 170 175

Ile Leu Gly Asp Gly Asn Ala Lys Val Val Tyr Val Lys Glu Glu Asp
180 185 190

Ile Gly Thr Phe Thr Ile Lys Ala Val Asp Asp Pro Arg Thr Leu Asn
195 200 205

Lys Thr Leu Tyr Leu Arg Leu Pro Ser Asn Thr Tyr Ser Phe Asn Glu
210 215 220

Leu Ala Ala Leu Trp Glu Thr Lys Ile Gly Lys Thr Leu Glu Lys Val
225 230 235 240

Tyr Val Pro Glu Glu Ala Val Leu Lys Lys Ile Glu Glu Thr Pro Phe
245 250 255

Pro Ala Ser Ile Gly Ile Ser Ile Gly Tyr Ser Thr Phe Val Lys Gly
260 265 270

Asp Gln Thr Asn Phe Glu Ile Gly Pro Asp Gly Val Glu Gly Ser Gln
275 280 285

Leu Tyr Pro Asp Val Lys Tyr Thr Thr Val Glu Glu Tyr Leu Asp Gln
290 295 300

Phe Val
305

<210> 7
<211> 392
<212> DNA
<213> Forsythia x intermedia

<400> 7
atcattgggg gtacagggta cttagggagg agattggtta aggcaagttt agctcaaggt 60

catgaaacat acattctgca taggcctgaa attggtgttg atattgataa agttgaaatg 120

ctaatatcat ttaaaatgca aggagctcat cttgtatctg gttctttcaa ggatttcaac 180

agtctggtcg aggctgtcaa gctcgtagac gtagtaatca gcgccatttc tggtgttcat 240

attcgaagcc atcaaattct tcttcaactc aagcttgttg aagctattaa agaggctgga 300

aatgtcaaga gatttttacc atctgagttt ggaatggatc ctgcaaaatt tatggatacg 360

gccatggaac ccggaaaggt aacacttgat ga 392

<210> 8
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide

<220>
<221> misc_feature
<222> (1)..(20)
<223> PCR primer PLRNT1 wherein n at positions 9, 12, 15
and 18 represents a, c, g or t

<400> 8 20
athathggng gnacnggnta

<210> 9
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide

<220>
<221> misc_feature
<222> (1)..(20)
<223> PCR primer PLRI5R wherein n at positions 6, 9, 12,
and 18 represents a, c, g or t.

<400> 9 20
tcytcnarng tnacyttncc

<210> 10
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide

<220>
<221> misc_feature
<222> (1)..(36)
<223> PCR primer PT-ATG-NdeI

<400> 10 36
ttcagggccc atatgggaag caggagcagg atactc

<210> 11
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:

oligonucleotide

<220>
<221> misc_feature
<222> (1)..(36)
<223> PCR primer PT-Rev-NdeI

<400> 11
tgtcgaatac atatgaaagg cgataaccaa caattt 36

<210> 12
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide

<220>
<221> misc_feature
<222> (1)..(20)
<223> PCR primer Cj-PCBERNT wherein n at positions 8 and 15
represent a, c, g or t

<400> 12
atwggmgnaa cwggntacat 20

<210> 13
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
oligonucleotide

<220>
<221> misc_feature
<222> (1)..(20)
<223> PCR primer Cj-PCBER-CT wherein n at positions 12
and 15 represent a, c, g or t

<400> 13
tayacmacyg tngangagta 20

INTERNATIONAL SEARCH REPORT

International application No.
PC US99/16746

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12N 9/02, 15/00; C07H 21/02, 21/04
US CL :435/189, 320.1, 468; 536/23.1, 23.2

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/189, 320.1, 468; 536/23.1, 23.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P --- Y,P	Gang, D. et al. Evolution of Plant Defense Mechanisms. J. Biol. Chem. 12 March 1999. Vol. 274, No. 11, pages 7516-7527, see entire document.	1-6, 13, 15-18, 23, 24, 26, 28, 30 --- 7-12, 14, 19-22,25,27,29,31-33

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
30 SEPTEMBER 1999

Date of mailing of the international search report
04 NOV 1999

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
PETER TUNG
Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/16746

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

West 1.2, MEDLINE, CANCERLIT, CAPLUS, BIOSIS, EMBASE
search terms: dehydrodiconiferyl alcohol benzylic ester, reductase, norman lewis, hiroyuki kasahara, david gang, laurence davin

BLANK PAGE